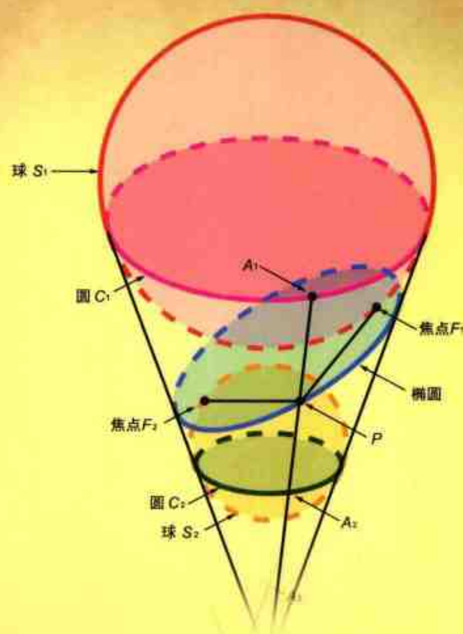
TURING 图灵数学·统计学丛书 50

WILEY



Linear Algebra

A First Course with Applications to Differential Equations

线性代数及其应用导论

[美] Tom M. Apostol 著

沈灏 沈佳辰 译



人民邮电出版社
POSTS & TELECOM PRESS

线性代数及其应用导论

Linear Algebra A First Course with Applications to Differential Equations

"一本用来学习线性代数、理解抽象向量空间的佳作，写作巧妙，通俗易懂。"

——亚马逊读者评论

本书是由美国著名数学教育家撰写的经典教材，不仅介绍了向量代数、线性空间、线性变换、矩阵、行列式和二次型等传统授课内容，还介绍了线性代数在微分方程中的应用。书中内容独具特色，自成体系，理论和应用并重。书中习题丰富，并且提供了习题解答，便于课堂教学或自学。

本书篇幅适中，叙述简洁，通俗易懂，是一本非常好的线性代数入门教材，已被很多学校采用。

Tom M. Apostol 加州理工学院荣休教授，著名的解析数论专家和数学教育家，美国数学学会和科学发展协会会士。1923年出生于美国犹他州，父母均为希腊移民。分别于1946年和1948年获得华盛顿大学西雅图分校硕士学位和加州大学伯克利分校博士学位，此后在加州大学伯克利分校和MIT任教，1950年加入加州理工学院。2001年当选雅典科学院通讯院士。Apostol教授著述颇丰，除本书外还著有《解析数论导引》、《微积分》（卷 I 和卷 II）以及《数学分析》等专著和教材，他的这些著作在国际上已产生重要影响。

延伸阅读

Lax 《线性代数及其应用(第2版)》
Lay 《线性代数及其应用(第3版•修订版)》
Axler 《线性代数应该这样学》
Sauer 《数值分析》
Atkinson, Han 《数值分析导论(第3版)》
Demmel 《应用数值线性代数》
Trefethen, Bau 《数值线性代数》
Golub, Van Loan 《矩阵计算(第3版)》
Horn, Johnson 《矩阵分析(英文版)》，卷1和卷2



WILEY
www.wiley.com

图灵网站：www.turingbook.com 热线：(010)51095186
反馈/投稿/推荐信箱：contact@turingbook.com
有奖勘误：debug@turingbook.com

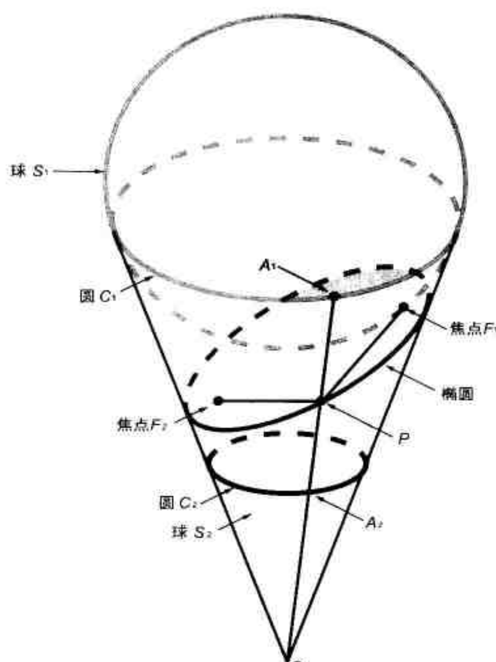
分类建议 数学/线性代数

人民邮电出版社网址：www.ptpress.com.cn



ISBN 978-7-115-23890-0
定价:59.00元

TURING 图灵数学·统计学丛书 50



Linear Algebra

A First Course with Applications to Differential Equations

线性代数及其应用导论

[美] Tom M. Apostol 著

沈灏 沈佳辰 译



人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目(CIP)数据

线性代数及其应用导论/(美)阿波斯托尔(Apostol, T.M.)著;沈灏,沈佳辰译. —北京:人民邮电出版社, 2010. 11

(图灵数学·统计学丛书)

书名原文:Linear Algebra: A First Course, with Applications to Differential Equations

ISBN 978-7-115-23890-0

Ⅰ.①线… Ⅱ.①阿… ②沈… ③沈… Ⅲ.①线性代数 Ⅳ.①O151.2

中国版本图书馆 CIP 数据核字(2010)第 182495 号

内 容 提 要

本书篇幅适中,叙述简洁,通俗易懂,是一本非常好的线性代数入门教材. 主要内容包括向量代数、线性空间、线性变换、矩阵、行列式和二次型等. 书中内容独具特色,自成体系,从基础知识讲起,进而进入线性代数的核心内容,最后达到应用,理论和应用并重. 最后 3 章介绍了线性代数在微分方程中的应用.

本书适合数学专业和其他各理工科专业高年级本科生和研究生作为教材,也值得数学教师和相关研究人员参考.

线性代数及其应用导论

◆ 著　[美]Tom M. Apostol
译　沈　灏　沈佳辰
责任编辑　明永玲

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
邮编 100061　电子函件 315@ptpress.com.cn
网址:http://www.ptpress.com.cn
北京铭成印刷有限公司印刷

◆ 开本:700×1000 1/16
印张:21.5
字数:431 千字　2010 年 11 月第 1 版
印数:1–3 000 册　2010 年 11 月北京第 1 次印刷

著作权合同登记号　图字: 01-2010-0828 号

ISBN 978-7-115-23890-0

定价:59.00 元

读者服务热线:(010)51095186　印装质量热线:(010) 67129223

反盗版热线:(010)67171154

版 权 声 明

译 者 序

本书是美国数学家 Tom M. Apostol 编著的著名教材, 是 Apostol 教授在其两卷本教材 *Calculus, Second Edition* 中线性代数部分的基础上整理补充和更新后单独组成的.

本书对线性代数的基本概念、基础理论、重要方法和技巧进行了比较深入系统的介绍. 特别是, 书中用了三章的篇幅论述线性代数在微分方程理论中的应用.

为了便于学生学习, 在正文之前对本书用到的一些预备知识做了简要的介绍.

本书由沈灏和沈佳辰翻译, 其中沈佳辰翻译本书前五章, 沈灏翻译其余各章.

最后, 对本书术语的翻译作一点说明, 即在一般的线性代数教材中, 通常也称抽象线性空间的元素为“向量 (vector)”, 这样比较有利于为抽象空间建立直观的几何类比. 但在本书中还是遵照英文原著的意思, 称抽象线性空间中的元素为“元素 (element)”. 不足之处, 敬请读者批评指证.

译 者 介 绍

沈　灏　1945 年生, 上海交通大学数学系教授、博士生导师, 研究方向: 组合设计, 有限几何与编码理论. 曾任中国数学会理事, 上海市数学会理事; 1993 年获 “全国优秀教师” 称号, 1998 年获 “宝钢优秀教师奖”, 2007 年获 “上海市教学名师奖” .

沈佳辰　1979 年生, 1991 年获上海交通大学学士学位, 2007 年获美国路易斯大学博士学位, 高级工程师, 研究方向: 计算机理论, 信息安全.

前　言

我的两卷本教材《微积分》包含了单变量与多变量函数的微积分、微分方程、无穷级数、线性代数、概率论与数值分析等内容. 多年来, 朋友们一直催促我在《微积分(第 2 版)》的基础上编写一本线性代数教材. 两卷本《微积分》已被翻译成意大利文并分成三卷出版, 其中第二卷介绍的就是线性代数内容, 从某种角度说, 朋友们对我的期望已经实现. 本书是作为与《微积分》的各卷彼此独立的教材编写的.

为安排有关背景材料的需要, 本书从回顾某些预备知识 (第 0 章) 开始. 预备知识分为两部分: 与微积分无关的预备知识, 用以理解从第 1 章到第 7 章的内容; 关于微积分的预备知识, 这会在第 8 章到第 10 章中用到. 第 1 章和第 2 章介绍 n 维空间中的向量代数及其在解析几何中的应用, 这两章为第 3 章到第 7 章中关于线性代数的较为抽象的处理提供动力并给出实例.

第 3 章讨论线性空间、子空间、线性无关性、基与维数、内积、正交性以及 Gram-Schmidt 正交化方法. 第 4 章介绍线性变换和矩阵及其在线性方程组中的应用. 第 5 章通过行列式的性质用公理化的方法介绍行列式, 该章对行列式的处理比我在《微积分》中的处理简单. 第 6 章讨论特征值与特征向量, 包括用以导出 Cayley-Hamilton 定理的三角化定理, 还简要介绍了 Jordan 标准型. 第 7 章在欧氏空间的框架下继续讨论特征值和特征向量, 及其在二次型和二次曲线中的应用.

在第 3 章到第 7 章中, 微积分概念仅偶尔在例题与习题中出现而且易于识别, 即使略去或以后再看也不影响教材的连续性. 本书的这一部分可用作不需要用到关于微积分的预备知识的线性代数基础课程的教材. 然而, 本书更适合于那些在某种程度上学习过诸如初等微积分或有限数学等课程的读者.

第 8~10 章当然需要读者具有微积分背景. 第 8 章将线性代数概念用于 n 阶线性微分方程, 特别是常系数线性微分方程. 第 9 章利用矩阵分析讨论微分方程组, 本章主要集中在由线性代数与矩阵分析的交互作用导出指数矩阵的性质. 第 10 章利用 Picard 的逐次逼近法处理微分方程组的存在性和唯一性问题, 我们也用收缩算子的语言对此展开论述.

尽管本书大部分内容都取自我的《微积分》一书, 但还是对某些主题作了改动或重新安排, 此外, 还添加了一些新内容与新习题.

本书可用作大学本科一年级或二年级的教材, 对于那些在多年前学过数学但未学过线性代数的人, 如果现在希望学习线性代数及其基本概念又不过分强调抽象性与形式化, 本书也非常适合.

Tom M. Apostol
于加州理工学院

目　录

第0章　预备知识

本章的第一部分简要介绍本书中需要用到的与微积分无关的若干预备知识：关于实数、直角坐标系、复数、数学归纳法的一些结果. 第二部分则介绍本书中用到的关于微积分的若干预备知识. 第 1 章和第 2 章将会介绍向量代数及其在解析几何中的应用, 这两章不会用到微积分中的任何知识, 而且这两章的内容为由第 3 章开始介绍的线性代数的抽象处理提供了动力和实例. 在第 3 章到第 7 章中, 微积分仅在某些例题和习题中偶尔出现, 微积分出现的地方将会清楚地标示出来, 而且即使跳过也完全不会影响对本书的阅读.

线性代数和微积分虽然是两门完全不同的学科, 但是线性代数的一些最重要的应用需要微积分中诸如积分、导数和无穷级数等概念. 熟悉一元微积分对理解这些应用, 特别是最后三章中介绍的微分方程中的应用, 毫无疑问是至关重要的. 同时, 线性代数的应用也使得对微分方程的理解显得更为自然.

I.　与微积分无关的预备知识

0.1　用直线上的点表示实数

实数可以用几何方法表示为直线上的点. 在直线上先取定两个点, 左边的点表示 0, 右边的点表示 1, 如图 0.1 所示, 这样的取法给定了这个表示的单位长度. 由欧氏几何的公理可知, 每一个实数在此直线上都有且仅有一个点与之对应. 相应地, 此直线上的每一个点都有且仅有一个实数与之对应. 我们通常称这条直线为**实直线** (real line) 或**实轴** (real axis), 称实数 x 所对应的点为点 x. 我们用记号 $\mathbb{R}$ 来表示实数全体所组成的集合.

若 $x < y$, 则点 x 在点 y 的左边, 如图 0.1 所示. 任意正实数 x 对应的点在零点右方 x 长度处, 任意负实数 x 对应的点在零点左方 $|x|$ 长度处.

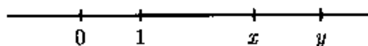


图 0.1　实数在直线上的几何表示

0.2　用平面上的点表示实数对

平面上的点可以用实数对来表示. 选定平面上两条相互垂直的参考直线：一条

水平的 x 轴和一条竖直的 y 轴. 两轴的交点称为**原点** (origin), 用 O 表示. 在 x 轴上取原点右方的一点表示 1, 它和原点之间的距离称为**单位距离** (unit distance). 通常在 y 轴上也使用相同的单位距离. 平面上的每一点都有一个实数对与之对应, 该实数对称为该点的 **坐标** (coordinate), 它告诉我们如何找到这个点. 如图 0.2 所示, 坐标为 $(3,2)$ 的点在 y 轴右方 3 个单位距离和 x 轴上方 2 个单位距离处, 其中数 3 称为该点的 **x 坐标或横坐标** (abscissa), 数 2 称为它的 **y 坐标或纵坐标** (ordinate). y 轴左方的点的横坐标为负, x 轴下方的点的纵坐标为负. 为了纪念解析几何奠基人之一 —— 勒内・笛卡儿 (Rene Descartes, 1596—1650), 将上述定义的点的坐标称为该点的**笛卡儿坐标** (Cartesian coordinate).

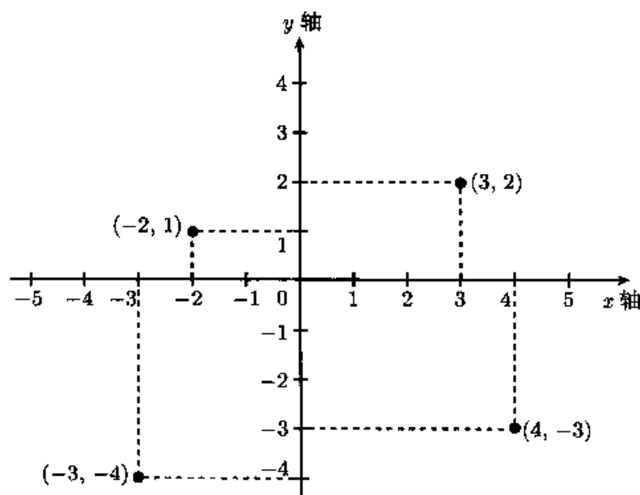


图 0.2 用实数对表示平面上的点

当我们用实数对表示一个点时, 一般将横坐标写在前面, 纵坐标写在后面. 因此, 我们使用的实数对 (a,b) 也称为有序对, 其中 a 称为它的**第一分量** (first entry), b 称为它的**第二分量** (second entry). 当且仅当 $a=c, b=d$ 时, 两个有序对 (a,b) 和 (c,d) 表示同一个点. 若 $a>0, b>0$, 则点 (a,b) 在第一象限中; 若 $a<0, b>0$, 则点 (a,b) 在第二象限中; 若 $a<0, b<0$, 则点 (a,b) 在第三象限中; 若 $a>0, b<0$, 则点 (a,b) 在第四象限中. 在图 0.2 中, 每个象限内均有一点.

类似地, 我们也可以确定空间中点的实数表示. 在空间中取三条交于一点 (原点) 且两两垂直的直线, 这三条直线决定了三个两两垂直的平面, 则空间中的每一个点可以由该点到这三个平面的有向距离唯一确定. 我们将在下一章中讨论三维笛卡儿坐标, 目前我们重点讨论二维的情形.

和平面中的曲线一样, 任何一个几何图形都是满足一个或多个条件的点的集合, 如果用坐标 x 和 y 来表示这些条件的话, 我们将得到一个或多个刻画该图形的

关系 (等式或不等式). 例如, 如图 0.3 所示, 考虑一个以原点为圆心, 以 r 为半径的圆. 用 (x,y) 表示此圆上的任意一点 P, 则线段 OP 是直角边长分别为 $|x|$ 和 $|y|$ 的直角三角形的斜边, 所以由勾股定理, 我们有

$$x^2+y^2=r^2.$$

该圆上的所有点都满足该方程, 且仅有该圆上的点满足该方程, 所以, 我们称它为该圆的**笛卡儿方程** (Cartesian equation). 该圆内的所有点满足不等式 $x^2+y^2<r^2$, 而圆外的所有点都满足不等式 $x^2+y^2>r^2$. 这个例子演示了解析几何如何用来将点的几何关系化为实数的代数关系.

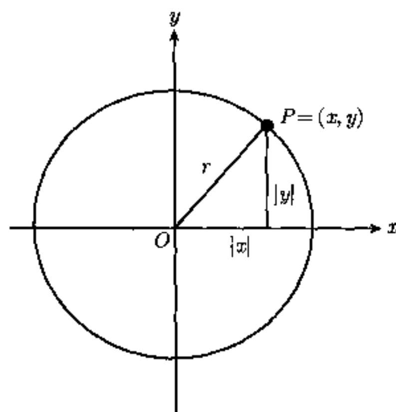


图 0.3 笛卡儿方程 $x^2+y^2=r^2$ 所表示的圆

0.3 极 坐 标

平面中的点也可以用极坐标来定位. 取 P 为平面上不同于原点的任意点, 设原点与点 P 之间的线段长为 $r>0$, 且该线段与正 x 轴之间的夹角为 θ, 如图 0.4 所示, 则 r 和 θ 称为 P 的**极坐标** (polar coordinate). 极坐标和直角坐标满足以下关系:

$$x=r\cos\theta,\ y=r\sin\theta. \tag{0.1}$$

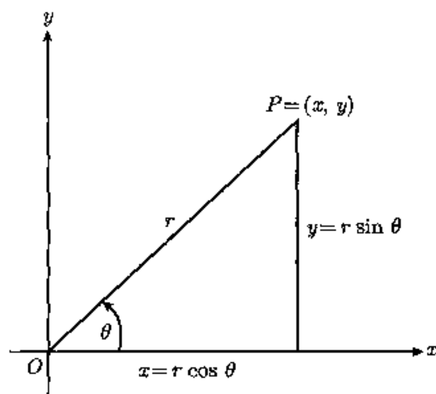


图 0.4 极坐标

我们称正数 r 为 P 的**极半径** (radial distance), 称 θ 为 P 的一个**极角** (polar angle). 考虑到若 θ 满足式 (0.1), 则对任意整数 n, $\theta+2n\pi$ 也满足式 (0.1), 所以我们称 θ 为 P 的一个极角而不是简单地称它为 P 的极角. 若实数对 (r,θ) 满足式 (0.1), 且 $r>0$, 则我们称它为 P 的极坐标.

极半径 r 可用公式 $r=\sqrt{x^2+y^2}$ 由 x 和 y 唯一确定, 不过极角 θ 不能由 x 和 y 唯一确定, 而是只能在允许相差 2π 的整数倍的意义下唯一确定.

若 P 是原点, 则 $r=0$ 和任意 θ 均满足关系式 (0.1), 所以我们用 $r=0$ 作为原点的极半径, 并用任意实数 θ 作为原点的极角.

用极坐标表示一些曲线比用直角坐标要方便得多. 例如, 圆心在原点, 半径为 2 的圆的笛卡儿方程为 $x^2+y^2=4$, 而在极坐标系中, 我们用更简单的方程 $r=2$ 来表示同一个圆, 而且圆内的点满足不等式 $r<2$, 圆外的点满足不等式 $r>2$.

0.4 复　　数

由于所有实数的平方都非负, 所以我们知道方程 $x^2+1=0$ 在实数系统内无解. 为了使此方程有解, 我们将引入一类新的数 —— 复数.

早在 16 世纪, 人们已经用符号 $\sqrt{-1}$ 来提供表示方程 $x^2+1=0$ 的解的方法, 这个符号后来用字母 i 代替, 它是一个虚构的、想象出来的数字, 同时除了平方是 -1 之外, 它也像所有的实数一样参与代数运算并满足这些运算的规则. 例如, 二次多项式可以进行如下因式分解:

$$x^2+1=x^2-\mathrm{i}^2=(x-\mathrm{i})(x+\mathrm{i}).$$

另外, 我们也用 $x=\pm\mathrm{i}$ 来表示方程 $x^2+1=0$ 的解, 而不管它的具体意义是什么. 我们称诸如 $2+3\mathrm{i}$ 的表达式为**复数** (complex number). 一开始, 人们仅仅是纯形式化地使用复数, 直到约 300 年后, 才开始使用符合现代标准的方式来定义它们.

在 19 世纪早期, Carl Friedrich Gauss (1777–1855) 和 William Rowan Hamilton (1805–1865) 几乎同时且相互独立地提出使用被赋予了某些特殊性质的有序实数对 (a,b) 来表示复数的思想. 现在复数的这种表示方法已经被广泛接受, 我们将在下一节中介绍这种思想.

0.5 复数的定义与代数性质

在 0.2 节中, 我们用实数对定义平面上点的直角坐标, 本节中, 为了对复数进行代数运算, 我们用满足下列条件的实数对定义复数.

定义　*设 a 和 b 是实数, 则 (a,b) 表示一个复数, 且两个复数之间的相等、加法、乘法关系由下式定义:*

(a) 相等: $(a,b)=(c,d)$ 当且仅当 $a=c$ 且 $b=d$.

(b) 加法: $(a,b)+(c,d)=(a+c,b+d)$.

(c) 乘法: $(a,b)(c,d)=(ac-bd,ad+bc)$.

相等的定义表明 (a,b) 是作为一个有序对使用的, 所以 $(2,3)$ 和 $(3,2)$ 是两个不同的复数. a 和 b 称为复数 (a,b) 的**分量**, 第一个分量 a 称作该复数的**实部** (real part), 第二个分量 b 称作该复数的**虚部** (imaginary part).

细心的读者会发现, 在这个定义中, 我们并没有使用 $\sqrt{-1}$ 这个符号, 事实上, 在本书中, 我们将把 i 看作一个特殊的复数, 它具有早先的数学家给虚构的符号 $\sqrt{-1}$

所赋予的所有代数性质, 为此, 我们先来讨论刚才定义的运算的基本性质.

定理 0.1 *复数的加法和乘法满足交换律、结合律、分配律. 即若 x, y, z 是任意复数, 则我们有如下性质:*

交换律: $x+y=y+x, xy=yx$,

结合律: $x+(y+z)=(x+y)+z, x(yz)=(xy)z$,

分配律: $x(y+z)=xy+xz$.

证明 由复数加法和乘法的定义, 易证本定理. 例如要证乘法结合律, 我们用分量形式表示复数 x, y, z, 设 $x=(x_1,x_2), y=(y_1,y_2), z=(z_1,z_2)$, 则有

$$\begin{aligned}x(yz)&=(x_1,x_2)(y_1z_1-y_2z_2,y_1z_2+y_2z_1)\\&=(x_1(y_1z_1-y_2z_2)-x_2(y_1z_2+y_2z_1),x_1(y_1z_2+y_2z_1)+x_2(y_1z_1-y_2z_2))\\&=((x_1y_1-x_2y_2)z_1-(x_1y_2+x_2y_1)z_2,(x_1y_2+x_2y_1)z_1+(x_1y_1-x_2y_2)z_2)\\&=(x_1y_1-x_2y_2,x_1y_2+x_2y_1)(z_1,z_2)\\&=(xy)z.\end{aligned}$$

交换律和分配律也可类似证明. □

我们可在复数中定义零、负数、倒数、商等基本代数概念, 它们的定义如下.

我们称复数 $(0,0)$ 为**复数零** (zero). 由于对任意复数 (a,b), 有 $(0,0)+(a,b)=(a,b)$, 所以复数零是复数加法的单位元. 类似地, 考虑到对任意复数 (a,b), 有 $(a,b)(1,0)=(a,b)$, 所以复数 $(1,0)$ 是复数乘法的单位元.

由于 $(-a,-b)+(a,b)=(0,0)$, 我们称复数 $(-a,-b)$ 为 (a,b) 的**负元** (negative), 记作 $-(a,b)$.

我们定义两个复数 (a,b) 和 (c,d) 的**差** (difference) $(a,b)-(c,d)$ 为 (a,b) 与 (c,d) 的负元的和, 即 $(a,b)-(c,d)=(a,b)+(-(c,d))$.

所有的非零复数 (a,b) 都有一个相对于乘法单位元 $(1,0)$ 的**倒数** (reciprocal), 记为 $(a,b)^{-1}$. $(a,b)^{-1}$ 的定义如下式:

$$(a,b)^{-1}=\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right), \tag{0.2}$$

其中 $(a,b)\neq(0,0)$. 这个定义满足 $(a,b)(a,b)^{-1}=(1,0)$. 因为 $(a,b)\neq(0,0)$, 则 $a^2+b^2\neq 0$, 所以 $(a,b)^{-1}$ 对任意 $(a,b)\neq(0,0)$ 有意义.

我们定义两个复数 (a,b) 和 (c,d) 的**商** (quotient)$(a,b)/(c,d)$ 为 (a,b) 与 (c,d) 的倒数 $(c,d)^{-1}$ 的积, 即 $(a,b)/(c,d)=(a,b)(c,d)^{-1}$, 其中 $(c,d)\neq(0,0)$.

0.6 复数作为实数的推广

我们用记号 $\mathbb{C}$ 表示所有复数的集合, 设 $\mathbb{C}_0$ 是所有形如 $(a,0)$ 的复数的集合, 则 $\mathbb{C}_0$ 是 $\mathbb{C}$ 的一个子集, 且它包含所有虚部为零的复数. $\mathbb{C}_0$ 中任意两个元素的和与积仍然是 $\mathbb{C}_0$ 中的元素. 事实上, 我们有

$$(a,0)+(b,0)=(a+b,0), \quad (a,0)(b,0)=(ab,0).$$

这两个式子表明当求 $\mathbb{C}_0$ 中两个元素的和 (或积) 时, 我们只需要求它们实部的和 (或积), 所以对于加法和乘法, $\mathbb{C}_0$ 中的数就像实数一样. 而且对于减法和除法也是如此, 这是因为 $-(a,0)=(-a,0)$, $(b,0)^{-1}=(b^{-1},0)$, 其中 $b\neq 0$. 正因如此, 实数 x 和复数 $(x,0)$ 并无本质上的区别, 我们认为 x 和 $(x,0)$ 是相同的, 并将 x 记为 $(x,0)$, 特别将 0, 1, -1 分别记为 $(0,0)$, $(1,0)$, $(-1,0)$. 这样, 我们可以将复数系统看作是实数系统的一个推广.

而这么做同样也有几何意义, 在 0.9 节中, 我们将用平面上笛卡儿坐标为 x, y 的点来表示复数 (x,y), 此时 x 轴上的点即为集合 $\mathbb{C}_0$ 的几何表示.

0.7 虚数单位 i

复数具有一些实数不具备的代数性质, 例如二次方程 $x^2+1=0$ 无实数解, 但是它可用复数求解. 事实上, 复数 $(0,1)$ 就是它的一个解, 这是因为

$$(0,1)^2=(0,1)(0,1)=(0\cdot 0-1\cdot 1, 0\cdot 1+1\cdot 0)=(-1,0)=-1,$$

定义 用符号 i 表示复数 $(0,1)$, 并称它为*虚数单位*.

虚数单位具有平方为 -1 的性质, 即 $\mathrm{i}^2=-1$, 从而 $x=\mathrm{i}$ 是二次方程 $x^2+1=0$ 的解. 读者容易验证 $x=-\mathrm{i}$ 也是此方程的解.

现在我们可以建立复数的有序实数对表示与早期数学家使用的记号之间的关系了. 首先, 由复数乘法定义可得 $(b,0)(0,1)=(0,b)$, 于是我们有

$$(a,b)=(a,0)+(0,b)=(a,0)+(b,0)(0,1).$$

因此, 若记 $a=(a,0)$, $b=(b,0)$, $\mathrm{i}=(0,1)$, 则有 $(a,b)=a+b\mathrm{i}$. 由此, 我们证明了如下定理.

定理 0.2 *任意复数 (a,b) 都可表示为 $(a,b)=a+b\mathrm{i}$ 的形式.*

这种记法使关于复数加法和乘法的运算更便捷. 例如, 计算 $a+b\mathrm{i}$ 和 $c+d\mathrm{i}$ 的积, 运用分配律和结合律以及 $\mathrm{i}^2=-1$, 有

$$(a+b\mathrm{i})(c+d\mathrm{i})=ac-bd+(ad+bc)\mathrm{i},$$

此式显然符合复数乘法的定义. 类似地, 计算非零复数 $a+b\mathrm{i}$ 的倒数, 我们有

$$\frac{1}{a+b\mathrm{i}}=\frac{a-b\mathrm{i}}{(a+b\mathrm{i})(a-b\mathrm{i})}=\frac{a-b\mathrm{i}}{a^2+b^2}=\frac{a}{a^2+b^2}-\frac{b\mathrm{i}}{a^2+b^2},$$

上式满足等式 (0.2).

运用复数, 不仅可以求解简单的二次方程 $x^2+1=0$, 也可以求解一般的二次方程 $ax^2+bx+c=0$, 其中 a, b, c 均为实数, 且 $a \neq 0$. 用配方法, 我们可以将该二次方程写作如下形式:

$$\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a^2}=0.$$

若 $4ac-b^2 \leqslant 0$, 则此方程有实根 $(-b \pm \sqrt{b^2-4ac})/(2a)$. 若 $4ac-b^2>0$, 且 x 为实数, 则上式的等号左边必为正, 故此方程无实根, 然而, 它却有如下式所示的两个复根:

$$r_1=-\frac{b}{2a}+\mathrm{i}\frac{\sqrt{4ac-b^2}}{2a}, \quad r_2=-\frac{b}{2a}-\mathrm{i}\frac{\sqrt{4ac-b^2}}{2a}. \tag{0.3}$$

在 1799 年, Gauss 证明了, 对于形如 $a_0+a_1x+a_2x^2+\cdots+a_nx^n=0$ 的多项式方程, 其中 $a_0, a_1, \cdots, a_n$ 为任意给定实数, 且 $a_n \neq 0$, 若 $n \geqslant 1$, 此方程必有复数解. 不仅如此, 即使 $a_0, a_1, \cdots, a_n$ 为复数, 此方程仍然有复数解. 我们将这个结论称为**代数基本定理** (fundamental theorem of algebra), 它告诉我们, 不需要为了求解复系数多项式方程而去构造比复数更为复杂的数.

0.8 习　　题

1. 若两个复数的积为零, 证明其中至少有一个复数为零.
2. 证明二次方程 $x^2+1=0$ 仅有 $x=\mathrm{i}$ 和 $x=-\mathrm{i}$ 两个根.
3. 如果我们用一个较简单的类似加法定义的等式 $(a,b)(c,d)=(ac,bd)$ 取代 0.5 节中给出的定义, 来定义复数乘法.
   (a) 证明这样定义的乘法满足交换律和结合律, 同时也满足分配律.
   (b) 给出你认为此定义不适合作为复数乘法定义的两个理由.

0.9 几何解释 · 模与辐角

由于复数 (x,y) 是一个有序实数对, 它可以用平面上的一个点来表示, 也可以用从原点到点 (x,y) 的向量来表示, 如图 0.5 所示. 因此, 我们通常把 xy 平面称为复平面, 称 x 轴为实轴, y 轴为虚轴. 我们也经常将复数和点混用, 所以我们常称复数 z 对应的点为点 z.

复数的加法和减法运算有比较简单的几何解释. 用从原点到点 z_1 和 z_2 的向量分别表示复数 z_1 和 z_2, 则可用**平行四边形法则** (parallelogram law) 决定它们的和 z_1+z_2. 如图 0.6 所示, 以 z_1 和 z_2 的向量为边做平行四边形, 则它的通过原点的对角线即为从原点到 z_1+z_2 的向量, 而且另一条对角线则与 z_1 和 z_2 的差有关,

从 z_1 到 z_2 的向量和从 O 到 $z_2 - z_1$ 的向量平行且长度相等, 而从 z_2 到 z_1 的向量也和从 O 到 $z_1 - z_2$ 的向量平行且长度相等.

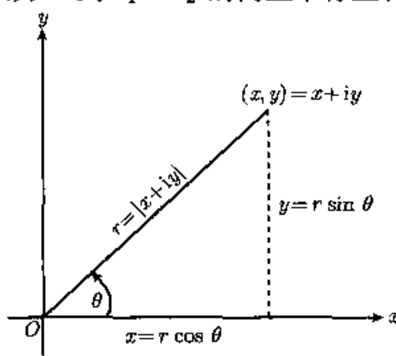


图 0.5　复数 $x + \mathrm{i}y$ 的几何解释

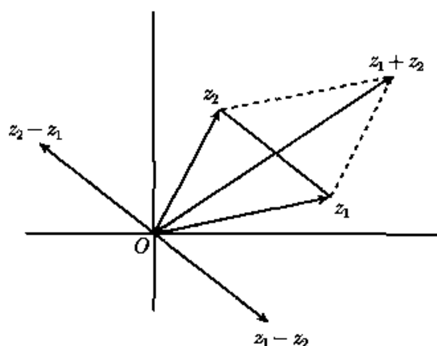


图 0.6　用平行四边形法则表示复数加法和减法的几何解释

若 $(x, y) \neq (0, 0)$, 则我们可以用极坐标来表示 x 和 y:

$$x = r\cos\theta, \quad y = r\sin\theta,$$

由此可知 (如图 0.5 所示)

$$x + \mathrm{i}y = r(\cos\theta + \mathrm{i}\sin\theta).$$

其中正实数 r 表示从原点到点 (x, y) 的距离, 我们将它称为 $(x + \mathrm{i}y)$ 的**模** (modulus) 或**绝对值** (absolute value), 并记作 $|x + \mathrm{i}y|$. 从而有

$$|x + \mathrm{i}y| = \sqrt{x^2 + y^2}.$$

我们称 (x, y) 的极角 θ 为 $x + \mathrm{i}y$ 的一个**辐角** (argument). 考虑到对任意整数 n, $\theta + 2n\pi$ 也是 $x + \mathrm{i}y$ 的辐角, 所以我们将 θ 称为 $x + \mathrm{i}y$ 的一个辐角而不称为 $x + \mathrm{i}y$ 的辐角. 有时我们希望对一个复数指定唯一的辐角, 这可通过规定辐角必须在一个给定的长为 2π 的半开半闭区间内取值即可. 通常我们使用 $[0, 2\pi)$ 或 $(-\pi, \pi]$ 作为辐角的取值区间. 特别地, 在本书中我们使用 $(-\pi, \pi]$, 如果复数 $x + \mathrm{i}y$ 的一个辐角落在此区间内, 则我们称它为复数 $x + \mathrm{i}y$ 的**主辐角** (principal argument), 并将它记作 $\arg(x + \mathrm{i}y)$, 即若 $x + \mathrm{i}y \neq 0$ 且 $r = |x + \mathrm{i}y|$, 我们定义 $\arg(x + \mathrm{i}y)$ 为满足以下条件的唯一实数 θ:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad -\pi < \theta \leqslant \pi.$$

我们认为复数零的模为 0, 且辐角可为任意实数 θ.

由于复数 z 的绝对值是一个线段的长度, 考察这些线段, 我们可以得到与实数绝对值同样的性质. 例如,

$$若\ z \neq 0\ 则\ |z| > 0, \quad |z_1 - z_2| = |z_2 - z_1|.$$

从几何意义上说, 绝对值 $|z_1 - z_2|$ 是复平面上点 z_1 和点 z_2 之间的距离.

0.10 共轭复数

设 $z = x + \mathrm{i}y$, 则 z 的**共轭复数** (complex conjugate) 为复数 $\overline{z} = x - \mathrm{i}y$. $\overline{z}$ 的几何意义为它与 z 关于实轴对称, 它们的实部相同, 而虚部大小相同, 符号相反. 共轭复数的定义表明 $|\overline{z}| = |z|$ 且

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}, \quad \overline{z_1/z_2} = \overline{z_1}/\overline{z_2}, \quad z\overline{z} = |z|^2.$$

由这些性质, 我们有 $|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1 \overline{z_1} z_2 \overline{z_2} = |z_1|^2 |z_2|^2$, 故

$$|z_1 z_2| = |z_1||z_2|. \tag{0.4}$$

同理, 若 $z_2 \neq 0$, 则有 $|z_1/z_2| = |z_1|/|z_2|$. 除此以外, 三角不等式也成立:

$$|z_1 + z_2| \leqslant |z_1| + |z_2|. \tag{0.5}$$

现在来证明不等式 (0.5). 我们知道

$$|z_1 + z_2|^2 = (z_1 + z_2)(\overline{z_1} + \overline{z_2}) = |z_1|^2 + |z_2|^2 + z_1\overline{z_2} + \overline{z_1} z_2,$$

由于一个复数和它的共轭复数的和为它的实部的 2 倍, 且一个复数的实部不会比它的模更大, 所以有

$$z_1\overline{z_2} + \overline{z_1} z_2 \leqslant 2|z_1\overline{z_2}| = 2|z_1||z_2|.$$

因此

$$|z_1 + z_2|^2 \leqslant |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = (|z_1| + |z_2|)^2,$$

从而可知三角不等式 (0.5) 成立.

若一个实系数二次方程无实根, 则由式 (0.3) 可求得它的两个复根, 且这两个复根互为共轭. 另一方面, 若复数 r_1 和 r_2 互为共轭, 令 $r_1 = \alpha + \mathrm{i}\beta$, $r_2 = \alpha - \mathrm{i}\beta$, 其中 α 和 β 为实数, 则 r_1 和 r_2 必为一个实系数二次方程的根. 事实上,

$$r_1 + r_2 = 2\alpha, \quad r_1 r_2 = \alpha^2 + \beta^2,$$

此外, 我们还有

$$(x - r_1)(x - r_2) = x^2 - (r_1 + r_2)x + r_1 r_2,$$

所以以 r_1 和 r_2 为根的实系数二次方程是

$$x^2 - 2\alpha x + \alpha^2 + \beta^2 = 0.$$

0.11 习　题

1. 将下列复数写成 $a + b\mathrm{i}$ 的形式.
   (a) $(1+\mathrm{i})^2$
   (b) $1/\mathrm{i}$
   (c) $1/(1+\mathrm{i})$
   (d) $(2+3\mathrm{i})(3-4\mathrm{i})$
   (e) $(1+\mathrm{i})/(1-2\mathrm{i})$
   (f) $\mathrm{i}^5 + \mathrm{i}^{16}$
   (g) $1 + \mathrm{i} + \mathrm{i}^2 + \mathrm{i}^3$
   (h) $\frac{1}{2}(1+\mathrm{i})(1+\mathrm{i}^{-8})$

2. 计算下列复数的绝对值.
   (a) $1+\mathrm{i}$　　(b) $3+4\mathrm{i}$
   (c) $(1+\mathrm{i})/(1-\mathrm{i})$　　(d) $1+\mathrm{i}+\mathrm{i}^2$
   (e) $\mathrm{i}^7+\mathrm{i}^{10}$　　(f) $2(1-\mathrm{i})+3(2+\mathrm{i})$
3. 计算下列复数的模和主辐角.
   (a) $2\mathrm{i}$　　(b) $-3\mathrm{i}$
   (c) -1　　(d) 1
   (e) $-3+\sqrt{3}\mathrm{i}$　　(f) $(1+\mathrm{i})/\sqrt{2}$
   (g) $(-1+\mathrm{i})^3$　　(h) $(-1-\mathrm{i})^3$
   (i) $1/(1+\mathrm{i})$　　(j) $1/(1+\mathrm{i})^2$
4. 分别求出满足下列各等式的实数 x 和 y.
   (a) $x+\mathrm{i}y=x-\mathrm{i}y$　　(b) $x+\mathrm{i}y=|x+\mathrm{i}y|$
   (c) $|x+\mathrm{i}y|=|x-\mathrm{i}y|$　　(d) $(x+\mathrm{i}y)^2=(x-\mathrm{i}y)^2$
   (e) $(x+\mathrm{i}y)/(x-\mathrm{i}y)=x-\mathrm{i}y$　　(f) $\sum\limits_{k=0}^{100}\mathrm{i}^k=x+\mathrm{i}y$
5. 分别画出满足下列各条件的所有复数 z 组成的集合的简图.
   (a) $|z|<1$　　(b) $z+\overline{z}=1$
   (c) $z-\overline{z}=\mathrm{i}$　　(d) $|z-1|=|z+1|$
   (e) $|z-\mathrm{i}|=|z+\mathrm{i}|$　　(f) $z+\overline{z}=|z|^2$
6. 令 f 为实系数多项式.
   (a) 证明 $\overline{f(z)}=f(\overline{z})$.
   (b) 由 (a) 证明: 若 f 有非实零点, 则 f 的所有非实零点必共轭成对出现.
7. 分别画出满足下列各条件的复数 z 组成的集合的简图.
   (a) $|2z+3|<1$　　(b) $|z+1|<|z-1|$
   (c) $|z-\mathrm{i}|\leqslant|z+\mathrm{i}|$　　(d) $|z|\leqslant|2z+1|$
8. 令 $w=(az+b)/(cz+d)$, 其中 a, b, c, d 为实数, 证明

$$w-\overline{w}=(ad-bc)(z-\overline{z})/|cz+d|^2.$$

   若 $ad-bc>0$, 证明 z 的虚部和 w 的虚部的正负号相同.

0.12　数学归纳法

我们知道没有最大的正整数, 这是因为对于任意整数 k, 总可以让它加 1 而得到整数 $k+1$, 显然 $k+1$ 大于 k. 所以说有无限多个正整数. 尽管如此, 我们总可以从 1 开始, 以从正整数 k 到 $k+1$ 为一步, 在连续的有限步内到达任意预先给定的正整数. 而这正是被数学家们称为归纳证明 (proof by induction) 的推理方法的基础.

归纳证明法　令 $A(n)$ 为与整数 n 有关的命题 (可能为真也可能为假), 如果

(a) $A(1)$ 为真,

(b) 对任意整数 $k\geqslant 1$, 由 $A(k)$ 为真可得 $A(k+1)$ 为真,

则 $A(n)$ 对所有正整数 $n\geqslant 1$ 都为真. □

可以用许多非数学的方法来阐述归纳法的基本思想. 例如, 假设有一列玩具兵向右无限延伸, 用连续的号码给这些玩具兵编号, 如图 0.7 所示. 假定它们的排列方式使得如果任意一个玩具兵向后摔倒, 它都会将它后面的那个玩具兵碰倒. 设前面那个玩具兵的编号为 k, 则后面的玩具兵的编号为 $k+1$, 如图 0.8 所示, 那么, 如果我们将编号为 1 的玩具兵碰倒的话, 将会发生什么事就非常明显了. 同样, 我们也容易知道如果碰倒编号为 n_1 的玩具兵的话, 那么它后面所有的玩具兵都会摔倒. 这种情况恰恰描述了归纳原理的一种变形, 这个变形可表述为: 若 $A(n_1)$ 为真, 且对任意整数 $k \geqslant n_1$, 由 $A(k)$ 为真可得 $A(k+1)$ 为真, 则对任意 $n \geqslant n_1$, $A(n)$ 为真.

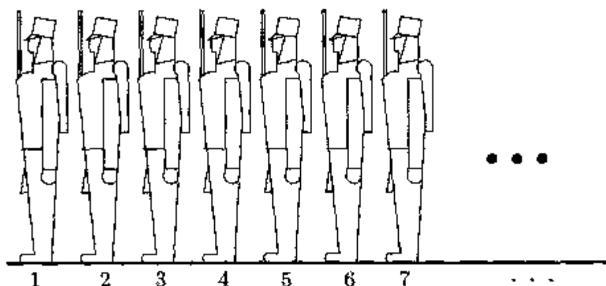


图 0.7

我们用下例来说明归纳证明法.

例 令 $A(n)$ 为命题 $(2n)!/(n!)^2 > 2^{n+2}$, 这个不等式当 $n = 1, 2, 3$ 时不成立, 但当 $n = 4$ 时成立. 我们可以用归纳法来证明对所有的 $n \geqslant 4$, 这个不等式都成立.

我们将会证明对所有的 $k \geqslant 1$, 由 $A(k)$ 为真可得 $A(k+1)$ 为真. 由命题 $A(k)$ 为真可知

$$\frac{(2k)!}{(k!)^2} > 2^{k+2}, \tag{0.6}$$

为了证明命题 $A(k+1)$ 为真, 只需证明

$$\frac{(2k+2)!}{((k+1)!)^2} > 2^{k+3}. \tag{0.7}$$

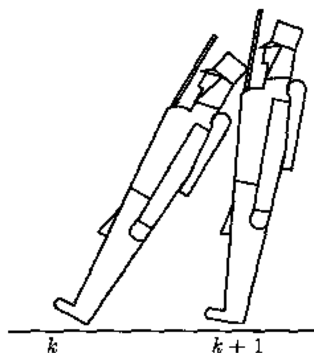


图 0.8

考察不等式 (0.7) 的左边:

$$\frac{(2k+2)!}{((k+1)!)^2} = \frac{(2k+2)(2k+1)(2k)!}{(k+1)^2(k!)^2} = \frac{2k+1}{k+1} \cdot \frac{2(2k)!}{(k!)^2},$$

由于 $2k+1 > k+1$, 所以 $(2k+2)!/((k+1)!)^2 > 2(2k)!/(k!)^2$, 而 $(2k)!/(k!)^2$ 即为不

等式 (0.6) 的左边, 故对所有的 $k \geqslant 1$, 若不等式 (0.6) 成立, 则不等式 (0.7) 也成立. 另外我们知道 $A(4)$ 为真, 故由归纳法得, 对所有 $n \geqslant 4$, 命题 $A(n)$ 为真. □

在上例中, 我们证明了对所有 $k \geqslant 1$, 若 $A(k)$ 为真, 则 $A(k+1)$ 为真, 即便当 $n = 1, 2, 3$ 时, $A(n)$ 为假, 但是由于 $A(4)$ 为真, 所以对 $n = 5, 6$ 及更大的整数, $A(n)$ 为真.

在本书中, 我们将会经常使用归纳证明法.

0.13 习　　题

1. 对所有整数 $n \geqslant 1$, 用归纳法证明下列公式:

   (a) $\sum\limits_{k=1}^{n} k = \frac{n(n+1)}{2}$,　　(b) $\sum\limits_{k=1}^{n} k^3 = \left(\sum\limits_{k=1}^{n} k\right)^2$.

2. 若一个实数 x 有性质 $x^2 = x + 1$, 用归纳法证明存在函数 $f(n)$, 使得对所有整数 $n \geqslant 1$, 都有 $x^{n+1} = f(n+1)x + f(n)$.

3. 令 $A(n)$ 表示命题 $\sum\limits_{r=1}^{n} r = \frac{(2n+1)^2}{8}$.

   (a) 证明对整数 $k \geqslant 1$, 由 $A(k)$ 可得 $A(k+1)$.

   (b) 试说明为什么命题 "由归纳法可知对所有 $n \geqslant 1$, $A(n)$ 为真" 不成立.

   (c) 将 $A(n)$ 中的等号改为一个不等号, 使得对所有 $n \geqslant 1$, $A(n)$ 都为真.

0.14 必要条件和充分条件

在本书中我们会经常讨论各种定理和它们的证明, 读者有必要理解这些讨论中出现的一些术语的意义.

一般地, 每一个定理都包含一个或多个形如 "由 A 可得 B" 的论断, 其中 A 为假设, B 为结论. 命题 "由 A 可得 B" 表示 "若 A 为真, 则 B 为真". 我们有时候也说 "A 是 B 的**充分条件** (sufficient condition)" 或 "B 是 A 的**必要条件** (necessary condition)". 例如, 平面几何中有一个简单的定理: **所有的等边三角形都是等腰三角形**. 这条定理可以表示为如下命题: **如果一个三角形是等边三角形, 那么它也是等腰三角形**. 这个定理具有 "由 A 可得 B" 的形式, 其中 A 表示命题 "一个三角形是等边三角形", B 表示命题 "一个三角形是等腰三角形".

如果一个定理具有 "由 A 可得 B" 的形式, 则它的逆命题为 "由 B 可得 A", 而这个命题可能为真也可能为假. 例如命题 "**所有的等边三角形都是等腰三角形**" 为真, 但是因为有一些等腰三角形并不等边, 所以 "**所有的等腰三角形都是等边三角形**" 为假.

如果定理 "由 A 可得 B" 及其逆命题 "由 B 可得 A" 均为真, 那么我们说 "A 是 B 的**充分必要条件** (necessary and sufficient condition)", 或者简单地说成 "当且

仅当 B 成立时 A 成立", 我们也称 A 和 B 是逻辑等价的.

为了证明两个命题 A 和 B 是逻辑等价的, 需要证明 "由 A 可得 B" 以及 "由 B 可得 A". 而为了证明三个命题 A, B, C 是逻辑等价的, 只需要证明由 A 可得 B, 由 B 可得 C, 由 C 可得 A, 这是因为由以上三条, 我们可以得到推论: 由 B 可得 A, 由 C 可得 B, 由 A 可得 C.

命题 B 的直接证明法是指对 B 的成立进行正面的逻辑论证. 而命题 B 的反证法则为: 假设 B 不成立 (即非 B 成立), 并由此得出某个命题 A 及其非同时成立的推断, 所以由这个自相矛盾的推断可知我们之前进行的假设不成立, 故 B 成立. 例如, 令 B 表示命题 "**没有最大的整数**", 我们用反证法证明 B 为真: 假设 B 为假, 则存在最大的整数, 设它为 n, 而 $n+1$ 也是整数且大于 n, 这与我们假定 n 是最大的整数矛盾, 所以 B 为真.

在本节的最后, 我们给出一个和上例有关的令人惊奇的定理, 读者可用反证法证明该定理.

定理 *如果存在最大正整数 n, 则 $n=1$.*

II. 关于微积分的预备知识

本书的最后三章将讨论线性代数在线性微分方程和微分方程组中的应用, 本章的第二部分复习在这些应用中需要用到的微积分知识. 我们将从微分开始.

0.15 导数概念

对一个定义在实轴的某一区间上的实值函数 f, 它的导数是对 f 进行**微分** (differentiation) 所得到的一个新函数, 这个过程可描述如下.

对 $h \neq 0$, 求得 $f(x+h)$ 和 $f(x)$ 这两个函数值的差, 然后用 h 除这个差值,

$$\frac{f(x+h)-f(x)}{h},$$

由此得到的值称为**差商** (difference quotient), 从几何上看, 它表示曲线 f 上点 $(x, f(x))$ 到点 $(x+h, f(x+h))$ 的弦的斜率. 我们也称它为从 x 到 $x+h$ 的区间上该函数的**平均变化率** (average rate of change).

令 h 趋近于零, 即令 h 的值越来越小, 我们观察在这个过程中差商的变化情况. 如果差商趋向于一个定值, 则称这个值为 f 在 x 上的**导数** (derivative), 并记作 $f'(x)$. 这个过程可用如下数学公式表示:

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

记号 $\lim\limits_{h\to 0}$ 读作 "当 h 趋向于零时的极限", 它表示导数 $f'(x)$ 是当 h 趋向于零时差商的极限值. 当该差商有极限值时, 称导数 $f'(x)$ **存在**, 并称函数 f 在 x **上可微** (differentiable). 导数也可记作

$$f'(x) = \mathrm{D}f(x) = \frac{\mathrm{d}}{\mathrm{d}x}f(x),$$

导数 $f'(x)$ 的几何意义为曲线 f 在点 $(x, f(x))$ 上的斜率, 而过点 $(x, f(x))$ 且以 $f'(x)$ 为斜率的直线称为曲线 f 在该点上的**切线**（tangent line）. 另外数 $f'(x)$ 也称作函数 f 在 x 处的**瞬时变化率**（instantaneous rate of change）.

导数是通过微分法由 f 得出的新函数, 而微分学是研究导数及其性质的一门学科.

0.16 导数的基本性质

由导数的定义可得以下基本性质, 其证明可在任何一本微积分教科书中找到, 这里就不再赘述了. 利用这些性质可简化导数的计算.

(1) **常数和函数之积的导数等于该常数与该函数的导数之积**　对任意常数c, 有

$$(cf)' = cf'.$$

(2) **两个函数之和的导数等于它们的导数之和**

$$(f+g)' = f' + g'.$$

将性质 (1) 和 (2) 可合并成一条性质, 称作导数的**线性性**:

(3) **对任意常数 a 和 b, 有**

$$(af + bg)' = af' + bg'.$$

以下两条性质可用来求函数的积和商的导数.

(4) **函数乘积的求导法则**　令 $f \cdot g$ 表示 f 和 g 的积, 则有

$$(f \cdot g)' = f \cdot g' + f' \cdot g.$$

(5) **函数商的求导法则**　对于 f 除以 g 的商, 在 $g(x) \neq 0$ 的点上我们有

$$\left(\frac{f}{g}\right)' = \frac{g \cdot f' - f \cdot g'}{g^2}.$$

下一个性质用来计算两个函数 f 和 g 的复合函数 $f \circ g$ 的导数, 其中

$$(f \circ g)(x) = f[g(x)].$$

(6) **复合函数的求导法则 (链式法则)**　令 h 为复合函数 $f \circ g$, 则有

$$h' = (f' \circ g) \cdot g'.$$

我们也将这个性质写为: 设 $h(x) = f[g(x)]$, 则

$$h'(x) = f'[g(x)] \cdot g'(x).$$

(7) **零导数定理** 若某个函数在一个开区间内的所有点上的导数均为零, 则该函数在此区间上的值为一个常数.

这个定理表明若两个函数在某个开区间上的导数相同, 则在此区间内, 它们的函数值的差为一个固定常数.

0.17 一些初等函数的导数

本书的最后三章将会用到一些初等函数的导数, 表 0.1 给出这些函数的导数, 其中设 n 和 k 为任意非零实常数.

表 0.1

函数 $f(x)$	导数 $f'(x)$
常数	零
x^n	nx^{n-1}
$\sin kx$	$k\cos kx$
$\cos kx$	$-k\sin kx$
e^{kx}	ke^{kx}

可用 0.16 节中的性质求上表中任意函数之和、积、商及复合函数的导数.

0.18 速度和加速度

导数在运动学的研究中有重要应用. 设一个物体在一条曲线上运动, 若在 t 时刻, 它到某一固定点的位移为 $f(t)$, 则导数 $f'(t)$ 称为该物体在 t 时刻的**速度**, 速度也是关于 t 的一个函数, 用 $v(t)$ 表示. 速度的导数 $v'(t)$ 称为该物体在 t 时刻的**加速度**. 加速度是位移的导数的导数, 也称为**二阶导数**, 记作 $f''(t)$. 例如, 若一个质点在 t 时刻的位移为

$$f(t) = \sin kt,$$

则该质点的速度为其位移的一阶导数

$$v(t) = f'(t) = k\cos kt,$$

它的加速度 $v'(t)$ 是其位移的二阶导数

$$f''(t) = -k^2\sin kt.$$

换言之, 位移函数 $f(t) = \sin kt$ 满足方程

$$f''(t) = -k^2 f(t).$$

这个方程称为简谐运动的微分方程. 简谐运动是指那些加速度与位移成正比, 但方向相反 (在方程中用负号表示) 的运动. 函数 $f(t) = \cos kt$ 也满足这个微分方程. 事实上, 这两个函数的所有形如

$$f(t) = a\cos kt + b\sin kt$$

的线性组合都满足这个微分方程, 其中 a 和 b 都是常数. 事实上, 人们还证明了所有满足简谐运动的微分方程的函数必是 $\sin kt$ 和 $\cos kt$ 的线性组合.

0.19 面积问题与积分学的历史

从本节开始, 我们复习积分学. 积分的历史可以追溯到古希腊时期. 从那时起的整整两千年里, 一个古老的问题一直向世界上的精英们挑战, 这就是求面积问题:

给定一个平面上的封闭曲线, 求一个正方形, 使它的面积和该曲线围成的区域的面积相同.

古希腊人试图用穷竭法来解决这个问题, 他们用边数越来越多的内接多边形来估计曲线围成区域的面积. 阿基米德 (Archimedes, 公元前 287— 前 212) 用这种方法求得了圆盘和其他一些特殊图形的面积, 这些特殊图形中包括了抛物线, 即求笛卡儿方程 $y = x^2$ 对应的抛物线与横轴上区间 $[0, t]$ 之间的区域的面积, 其中 $t > 0$. 这个区域在一个底边长为 t, 高为 t^2 的直角三角形内, 所以要求的面积必小于此直角三角形的面积 $\frac{1}{2}t^3$, 而阿基米德证明了它的面积恰好等于 $\frac{1}{3}t^3$.

在这之后, 求面积问题一直都没有什么进展. 直到几乎一千八百年后, 广泛使用精心挑选的代数符号重新激起了人们研究古老的穷竭法的兴趣. 在 16 世纪和 17 世纪, 诸如 Cavalieri, Toricelli, Roberval, Fermat, Pascal, Wallis 的先驱者们发现了很多关于求面积问题的零碎结果. 在 17 世纪之前, 穷竭法逐渐地变成了今天人们所熟知的积分法, 积分法是用于计算由曲线围成的平面图形的面积和由曲面围成的立体图形的体积的一种系统方法.

牛顿 (Isaac Newton, 1642—1727) 和莱布尼茨 (Gottfried Wilhelm Leibniz, 1646—1716) 两人是公认的积分学的发现者, 他们最主要的贡献在于统一了前人的工作成果并建立了积分和微分之间的联系.

正如导数是微分学最基本的概念一样, **积分** (integral) 是积分学中最基本的概念. 阿基米德关于抛物线面积的结果用积分的语言可以描述为:

$$x^2 \text{ 的从 } 0 \text{ 到 } t \text{ 的积分为 } \frac{1}{3}t^3.$$

用代数符号可以记为

$$\int_0^t x^2 \mathrm{d}x = \frac{1}{3}t^3.$$

记号 $\int$(拉伸的字母 S, 而 S 是 “和” 的拉丁语单词的首字母) 称为**积分号** (integral sign), 它是由 Leibniz 在 1675 年引入的. 这个记号提醒我们, 抛物线面积的计算是通过求许多高为 x^2、宽为 $\mathrm{d}x$ 的细长矩形的面积之和得到的. 记号 $\mathrm{d}x$ 表示 x 的微小改变, 而这个计算出 $\frac{1}{3}t^3$ 的过程称为**求积** (integration). 在积分号旁边的数字 0 和 t 称为**积分限** (limit). 我们必须特别指出, $\int_0^t x^2 \mathrm{d}x$ 是一个整体, 就好像字典里描述 lapidate 一词时不会把它分成 lap, id 和 ate 一样.

0.20 用积分法构造新函数

尽管积分法是因为计算面积而发明的, 但是它的作用决不仅仅在于计算面积, 它也可以用来计算弧长、体积、功、平均值、概率等量. 考虑到本书不是一本介绍微积分的书, 我们不给出积分的定义, 而只是将它看作由已知函数生成新函数的过程. 假设有一个 0.17 节表 0.1 中的初等函数 $f(x)$, 则它的积分 $\int_a^t f(x)\mathrm{d}x$ 是一个关于积分上限 t 的新函数, 我们用 $A(t)$ 表示,

$$A(t) = \int_a^t f(x)\mathrm{d}x.$$

其中 $f(x)$ 称为**被积函数** (integrand). 若被积函数 $f(x)$ 在区间 $[a,t]$ 上非负, 那么我们可以将 $A(t)$ 看作 $f(x)$ 的**面积函数** (area function), 该函数在 t 处的值等于曲线 $y = f(x)$ 和区间 $[a,t]$ 之间的区域的面积.

表 0.2 列出了本书最后三章中将要用到的一些初等函数的积分, 对于被积函数为非负的任意区间, 以及使被积函数连续的任意区间 $[0,t]$, 这些积分公式都成立.

表 0.2

被积函数 $f(x)$	积分 $A(t) = \int_0^t f(x)\mathrm{d}x$
常数 c	ct
x^n	$t^{n+1}/(n+1)$, 若 $n \neq -1$
$\sin kx$	$(1-\cos kt)/k$, 若 $k \neq 0$
$\cos kx$	$(\sin kt)/k$, 若 $k \neq 0$
e^{kx}	$(\mathrm{e}^{kt}-1)/k$, 若 $k \neq 0$
$(1+x)^{-1}$	$\ln(1+t)$, 若 $t > -1$

0.21 积分的基本性质

积分的以下基本性质可以用来简化积分的计算.

(1) $\int_u^t f(x)\mathrm{d}x = \int_a^t f(x)\mathrm{d}x - \int_a^u f(x)\mathrm{d}x$.

(2) $\int_u^t cf(x)\mathrm{d}x = c\int_u^t f(x)\mathrm{d}x$, c 为任意常数.

(3) $\int_u^t \{f(x)+g(x)\}\mathrm{d}x = \int_u^t f(x)\mathrm{d}x + \int_u^t g(x)\mathrm{d}x$.

性质 (2) 和 (3) 可以合并为一条性质, 叫作**线性性** (linearity), 即函数的线性组合的积分等于这些函数的积分的相同的线性组合:

(4) $\int_u^t \{af(x)+bg(x)\}\mathrm{d}x = a\int_u^t f(x)\mathrm{d}x + b\int_u^t g(x)\mathrm{d}x$.

(5) **比较性质** 若 $a < b$, 且在区间 $[a,b]$ 上都有 $f(x) \leqslant g(x)$, 则 $\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b g(x)\mathrm{d}x$.

这意味着 $|\int_a^b f(x)\mathrm{d}x| \leqslant \int_a^b |f(x)|\mathrm{d}x$.

以下两个定理建立了积分和微分之间的联系.

(6) **微积分第一基本定理**

若 $A(t)=\int_a^t f(x)\mathrm{d}x$, 其中 a 是一个常数, 则在 f 的所有连续点上导数 $A'(t)$ 存在且等于 $f(t)$.

换言之, 若被积函数连续, 则它的积分作为积分上限 t 的函数可导, 且它的导数 $A'(t)$ 等于被积函数在此上限上的值.

(7) **微积分第二基本定理**

若一个连续的被积函数 $f(x)$ 是某个函数 $P(x)$ 的导数, 则积分 $\int_a^t f(x)\mathrm{d}x$ 等于 $P(x)$ 在上限 t 和下限 a 上的值之差 $P(t)-P(a)$.

换言之, 若被积函数是另一个函数的导数, 则积分运算简化为减法运算:

$$\int_a^t P'(x)\mathrm{d}x = P(t)-P(a).$$

这个重要定理将积分的计算简化为寻找一个函数, 使它的导数恰好为被积函数. 上式也可写为如下形式:

$$P(t)=P(a)+\int_a^t P'(x)\mathrm{d}x,$$

由此式可知, 函数由它在一个定点上的值 $P(a)$ 和它在一个区间上的导数完全决定. 特别地, 若它的导数在 a 和 t 之间的区间上满足 $P'(x)=0$, 则 $P(t)=P(a)$. 换言之, 若一个函数在某个区间上的导数恒为零, 则它在该区间上是常数. 这就是 0.16 节中所提到的零导数定理.

对乘积 $f\cdot g$ 的导数用微积分第二基本定理, 有

$$\int_a^t (f\cdot g)'(x)\mathrm{d}x = f(t)g(t)-f(a)g(a),$$

因为我们有函数积的求导公式 $(f\cdot g)'=f\cdot g'+f'\cdot g$, 因此上式可表示为下述形式:

(8) **分部积分公式** $\int_a^t f(x)g'(x)\mathrm{d}x = f(t)g(t)-f(a)g(a)-\int_a^t g(x)f'(x)\mathrm{d}x.$

0.22 指数函数

本书第 8 章和第 9 章的内容与指数函数 e^x 密切相关, 读者需要了解指数函数 e^x. 指数函数常被称为最重要的实值初等函数. 对于实数 x, 我们有几种方式来定义 e^x, 其中之一是先用积分定义自然对数: 对于 $y>0$,

$$\ln y=\int_1^y \frac{1}{x}\mathrm{d}x,$$

然后定义指数函数为对数函数的逆函数, 即由 $x=\ln y$ 来定义 $y=\mathrm{e}^x$. 而另一种方式用幂级数来定义 e^x:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots.$$

我们知道, 上式对所有实数 x 收敛.

所有这些定义指数函数的方式都是等价的, 即由它们所定义的指数函数具有相同的基本性质. 本节简要地介绍这些性质, 下一节将把指数函数推广到复数域上, 使 e^z 对所有复数 z 都有意义. 在第 9 章中, 指数函数的概念甚至推广到指数矩阵.

下述为实指数函数的基本性质.

(1) $e^0 = 1$.

(2) **指数律** (law of exponents) 对所有实数 x 和 y, $e^x e^y = e^{x+y}$.

(3) **求导公式** 若 $f(x) = e^x$, 则 $f'(x) = e^x$.

即指数函数的导数与自己相同. 满足方程 $f'(x) = f(x)$ 的最普遍的函数就是 $f(x) = ce^x$, 其中 c 为常数.

(4) **恒正性**(positivity) 对所有实数 x, $e^x > 0$.

(5) **与自然对数的关系** 若 $y > 0$, 则当且仅当 $x = \ln y$ 时, $y = e^x$.

(6) **幂级数表示** 对所有实数 x, 有

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

0.23 复 指 数

我们可以对指数进行推广, 从而使得对任意给定复数 z, e^z 都有意义.

定义 *令 $z = x + iy$, 我们定义 e^z 为下式给出的复数:*

$$e^z = e^x(\cos y + i \sin y).$$

注意到当 $y = 0$ 时, $e^z = e^x$, 所以当 z 为实数时, e^z 和实指数函数的定义一致. 由定义也可推出指数律.

定理 0.3 *给定任意复数 a 和 b, 有*

$$e^a e^b = e^{a+b}. \tag{0.8}$$

证明 设 $a = x + iy$, $b = u + iv$, 则

$$e^a = e^x(\cos y + i \sin y), \quad e^b = e^u(\cos v + i \sin v),$$

它们的积为

$$e^a e^b = e^x e^u[(\cos y \cos v - \sin y \sin v) + i(\cos y \sin v + \sin y \cos v)].$$

由三角函数的加法公式和实指数函数的指数律有

$$e^a e^b = e^{x+u}[\cos(y + v) + i \sin(y + v)],$$

故式 (0.8) 得证. □

复指数函数具有一些实指数函数没有的性质. 例如, 若 x 为实数, 则当且仅当 $x=0$ 时, $\mathrm{e}^x=1$. 但是对于复数, 若对 $z=x+\mathrm{i}y$, $\mathrm{e}^z=1$, 则

$$\mathrm{e}^x\cos y=1,\quad \mathrm{e}^x\sin y=0.$$

因为 e^x 不可能为零, 故当且仅当 $\sin y=0$ 时, $\mathrm{e}^x\sin y=0$, 所以 $y=2n\pi$, 其中 n 为整数. 即当且仅当 $z=2n\pi\mathrm{i}$ 时,

$$\mathrm{e}^z=1,$$

其中 n 为整数.

实指数总是正的, 但是复指数可能为负, 例如 $\mathrm{e}^{\pi\mathrm{i}}=\cos\pi+\mathrm{i}\sin\pi=-1$.

由于正弦函数和余弦函数是以 2π 为周期的周期函数, 所以复指数函数的周期为 $2\pi\mathrm{i}$, 即 $\mathrm{e}^{z+2\pi\mathrm{i}}=\mathrm{e}^z$. 事实上, 因为 $\mathrm{e}^{2\pi\mathrm{i}}=1$, 由指数律即得 $\mathrm{e}^{z+2\pi\mathrm{i}}=\mathrm{e}^z\mathrm{e}^{2\pi\mathrm{i}}=\mathrm{e}^z$.

0.24　复数的极坐标形式

在 0.9 节中, 我们知道所有的非零复数 $z=x+\mathrm{i}y$ 都可以用极坐标的形式表示:

$$z=x+\mathrm{i}y=r(\cos\theta+i\sin\theta),$$

其中 $r=|z|$, $\theta=\arg z$. 用复指数, 我们可以将它进一步简化为

$$z=r\mathrm{e}^{\mathrm{i}\theta},$$

我们将它称为 z 的**极坐标形式** (polar form). 因为复指数的周期为 $2\pi\mathrm{i}$, 当将角 θ 替换为 $\arg z$ 与 2π 的任意整数倍之和时, 此等式仍然成立.

当我们进行复数的乘法时, 极坐标形式特别有用. 例如, 当我们计算两个用极坐标形式表示的复数 z 和 w 时, 设

$$z=r\mathrm{e}^{\mathrm{i}\theta},\quad w=R\mathrm{e}^{\mathrm{i}\phi},$$

则它们的积为

$$zw=(r\mathrm{e}^{\mathrm{i}\theta})(R\mathrm{e}^{\mathrm{i}\phi})=rR\mathrm{e}^{\mathrm{i}(\theta+\phi)}.$$

也就是说, 当两个复数相乘, 我们将它们的模相乘, 辐角相加.

复指数也可以用来求三角公式. 例如, 设 z 的模为 1, 则 $z=\mathrm{e}^{\mathrm{i}\theta}=\cos\theta+\mathrm{i}\sin\theta$. 从而对任意正整数 n, 都有

$$z^n=\mathrm{e}^{\mathrm{i}n\theta}=\cos n\theta+\mathrm{i}\sin n\theta,$$

即

$$(\cos\theta+\mathrm{i}\sin\theta)^n=\cos n\theta+\mathrm{i}\sin n\theta.$$

为了纪念数学家 Abraham de Moivre (1667—1754), 我们将这个公式称为**de Moivre 公式**. 用该公式, 我们可以简单推导关于正弦和余弦的多倍角公式. 例如, 当 $n=3$ 时, de Moivre 公式可以写成

$$(\cos\theta+\mathrm{i}\sin\theta)^3=\cos 3\theta+\mathrm{i}\sin 3\theta,$$

用二项式公式 $(a+b)^3=a^3+3a^2b+3ab^2+b^3$ 展开上面等式的左边, 其中令 $a=\cos\theta$, $b=\mathrm{i}\sin\theta$, 然后将实部和虚部分开, 我们有

$$(\cos\theta+\mathrm{i}\sin\theta)^3=\cos^3\theta-3\cos\theta\sin^2\theta+\mathrm{i}(3\cos^2\theta\sin\theta-\sin^3\theta).$$

因为两个复数相等的充要条件是它们的实部和虚部分别相等, 所以得

$$\cos 3\theta=\cos^3\theta-3\cos\theta\sin^2\theta,\quad \sin 3\theta=3\cos^2\theta\sin\theta-\sin^3\theta.$$

由 Pythagoras 恒等式 $\sin^2\theta+\cos^2\theta=1$, 我们用 $1-\cos^2\theta$ 替换上式中所有的 $\sin^2\theta$, 则有

$$\cos 3\theta=4\cos^3\theta-3\cos\theta,\quad \sin 3\theta=\sin\theta(4\cos^2\theta-1).$$

这个结论告诉我们 $\cos 3\theta$ 是 $\cos\theta$ 的一个三次多项式, 而 $\sin 3\theta$ 是 $\sin\theta$ 乘上一个 $\cos\theta$ 的二次多项式. 类似地, 我们也可证明, 对任意整数 $n>1$, $\cos n\theta$ 是 $\cos\theta$ 的 n 次多项式, 而 $\sin n\theta$ 是 $\sin\theta$ 乘以一个 $\cos\theta$ 的 $n-1$ 次多项式, 具体的公式可通过用二项式定理把 $(\cos\theta+\mathrm{i}\sin\theta)^n$ 展开, 并用 $1-\cos^2\theta$ 替换所有的 $\sin^2\theta$ 得到.

0.25 幂级数和函数级数

一个形如 $c_0+c_1x+c_2x^2+\cdots$ 的无穷级数称为关于 x 的**幂级数** (power series). 我们也将这种形式简洁地写成 $\sum\limits_{k=0}^{\infty}c_kx^k$, 其中 c_k 称作它的系数. 若 x 和所有的系数 c_k 都是实数, 则存在一个开区间 $(-r,r)$, 使得当 $|x|<r$ 时, 该级数收敛, 当 $|x|>r$ 时, 该级数发散, 这个区间称为此级数的**收敛区间** (interval of convergence). r 可以为零, 此时该级数仅当 $x=0$ 时收敛; r 也可为无穷大, 此时该级数对所有的实数收敛.

在收敛区间内, 级数的和定义了一个关于 x 的函数:

$$f(x)=\sum_{k=0}^{\infty}c_kx^k=\lim_{n\to\infty}\sum_{k=0}^{n}c_kx^k.$$

这个函数具有关于多项式的许多性质, 例如, 导数 $f'(x)$ 存在, 并可由对级数逐项求导而得:

$$f'(x)=\sum_{k=1}^{\infty}kc_kx^{k-1}.$$

若 $[a,b]$ 是 $f(x)$ 的收敛区间的一个子区间, 则积分 $\int_a^b f(x)\mathrm{d}x$ 可由对级数逐项求积而得:

$$\int_a^b f(x)\mathrm{d}x=\sum_{k=0}^{\infty}c_k\int_a^b x^k\mathrm{d}x.$$

在本书以后的章节中我们会遇到两个重要的级数: **几何级数** (geometric series)

$$\frac{1}{1-x}=\sum_{k=0}^{\infty}x^k,$$

与指数级数 (exponential series)

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

我们知道几何级数对 $|x|<1$ 收敛, 指数级数对所有实数 x 都收敛.

在第 10 章中, 我们还会见到更一般的在某个区间上具有**一致收敛性** (uniform convergence) 的形如 $\sum\limits_{k=0}^{\infty}u_k(x)$ 的级数. 级数在某个区间 I 上一致收敛的一个充分条件为, 该级数的所有项的绝对值在该区间上以一个收敛的正常数级数的对应项为上界, 即对 I 中所有的 x, 有 $|u_k(x)|\leqslant M_k$, 其中所有 M_k 都是正常数, 且 $\sum\limits_{k=0}^{\infty}M_k$ 收敛. 以下两条关于一致收敛级数的重要性质在第 10 章中将会用到.

(a) 若每一个函数 $u_k(x)$ 在 I 中的点 x_0 上都连续, 则和函数

$$f(x)=\sum_{k=0}^{\infty}u_k(x)=\lim_{n\to\infty}\sum_{k=0}^{n}u_k(x)$$

在 x_0 上也连续, 且 $f(x)$ 在 $x\to x_0$ 时的极限可由对级数的所有项求极限而得:

$$\lim_{x\to x_0}\sum_{k=0}^{\infty}u_k(x)=\sum_{k=0}^{\infty}\lim_{x\to x_0}u_k(x)=\sum_{k=0}^{\infty}u_k(x_0).$$

(b) 如果每一个函数 $u_k(x)$ 在区间 $[a,b]$ 上都连续, 则和函数的积分可由对级数逐项求积而得:

$$\int_a^b f(x)\mathrm{d}x=\sum_{k=0}^{\infty}\int_a^b u_k(x)\mathrm{d}x.$$

0.26 习 题

1. 令 $f(x)=\frac{1}{3}x^3-2x^2+3x+1$, 求 f 的图像上的所有点, 使得 f 在这些点上的切线平行于 x 轴.
2. 令 $f(x)=x^2+ax+b$, 其中 a 和 b 为常数, 求所有的 a 和 b 的值, 使得直线 $y=2x$ 是 f 在点 $(2,4)$ 上的切线.
3. 求所有 a, b, c 的值, 使得 $f(x)=x^2+ax+b$ 和 $g(x)=x^3-c$ 在点 $(1,2)$ 相交且它们在该点的切线相同.
4. 飞标发射之后 t 秒时距离地面高度为 $f(t)=144t-16t^2$, 其中 $0\leqslant t\leqslant 9$.
   (a) 画出 f 的图像.
   (b) 将飞标的速度表示成时间 t 的函数.
   (c) 证明当飞标到达最高点时其速度为零.
   (d) 证明飞标的加速度为常数.
5. 设 $x>0$, 令 $f(x)=(\ln x)/x$.

(a) 画出 f 在区间 $0 < x < \mathrm{e}^2$ 上的简图, 并分别描述满足 $f'(x) > 0$ 的点 x 和 $f''(x) > 0$ 的点 x.

(b) 求常数 c, 使得对所有 $t > 0$, $\int_1^t f(x)\mathrm{d}x = c(\ln t)^2$.

(c) 设直线 $y = mx$ 是 f 在点 $(a, f(a))$ 处的切线, 用 e 表示 m 及点 $(a, f(a))$ 的值.

(d) 用 e 表示实数 $b > 1$, 使得 f 平分以点 $(0,0)$, $(b,0)$, $(b, f(b))$ 为顶点的三角形的面积.

6. 令 $f(x) = 3 + \int_0^x \mathrm{e}^t/(1+t)\mathrm{d}t$, 不直接求此积分, 求一个三次多项式 $p(x)$, 使得 $p(0) = f(0)$, $p'(0) = f'(0)$, $p''(0) = f''(0)$, $p'''(0) = f'''(0)$.

7. 函数 f 对所有实数 x, y 满足函数方程

$$f(x+y) = f(x) + f(y) + xy.$$

(a) 证明 $f(0) = 0$.

(b) 令 $h \neq 0$, 证明

$$\frac{f(x+h) - f(x)}{h} = \frac{f(h) - f(0)}{h} + x,$$

并用此等式证明若 $f'(0)$ 存在, 则对所有的 x, $f'(x)$ 都存在, 且 $f'(x) = f'(0) + x$.

8. 令 $f(x) = \int_0^x g(t)\cos(x-t)\mathrm{d}t$, 其中 g 是一个给定的在任意点都可微的函数.

(a) 证明 $f'(x) = g(x) - \int_0^x g(t)\sin(x-t)\mathrm{d}t$.

[提示: $\cos(a-b) = \cos a\cos b + \sin a \sin b$.]

(b) 求满足 $f''(x) + f(x) = ag'(x) + bg(x)$ 的常数 a 和 b.

9. 用 0.24 节介绍的 de Moivre 公式证明以下各公式.

(a) $\cos 4x = 8\cos^4 x - 8\cos^2 x + 1$.

(b) $\sin 4x = \sin x(8\cos^3 x - 4\cos x)$.

(c) 证明 (a) 和 (b) 可通过微分相互推导.

第1章 向量代数

1.1 历史背景

早在古希腊时期, 人们就已经开始用数来表示直线上的点了. 到了 17 世纪, 笛卡儿进一步拓展了这一想法, 他用数对来表示平面上的点, 用数的三元组来表示空间中的点. 用这种方法, 他建立了解析几何. 在解析几何中, 我们用坐标的代数关系表示几何图形的性质.

贯穿其历史, 解析几何和微积分一直以来都是密不可分的, 它们之中一个领域取得的进展往往会促进另一个领域的发展. 在曲线上画切线的问题促成了微分学的发展, 而平面上求曲线所包围的面积问题又促成了积分学的发展.

随着这些成就的取得, 人们同时也获得了力学和数学物理方面的进展. 1788 年 Joseph-Louis Lagrange(1736—1813) 出版了他的著作《分析力学》(*Méchanique analytique*), 在这本书中, 他向人们展示了解析方法为力学研究带来的灵活性和巨大的作用. 在 19 世纪, 爱尔兰数学家 William Rowan Hamilton 引入了**四元数** (quaternion) 理论, 这种理论在当时提供了一种理解代数和物理的新方法和新视角. 主要是通过 J. W. Gibbs(1839—1903) 和 O. Heaviside(1850—1925) 的努力, 四元数分析和笛卡儿几何的优良特性得以结合, 并由此催生了一门称为**向量代数** (vector algebra) 的新学科. 人们也很快就认识到向量是研究和简化几何和物理中许多重要思想的理想工具. 本章讨论向量代数的基本概念和基础理论, 第 2 章将讨论向量代数在解析几何中的应用.

可以用三种不同的方法引进向量代数: 几何方法、解析方法和公理化方法. 若用几何方法, 我们用有向线段或箭头来表示向量. 关于向量的加法, 减法以及实数与向量的乘法等代数运算也将通过几何方法加以定义和研究.

若用解析方法, 我们将只用数来表示向量和向量运算, 这些数称为分量. 向量运算的性质则可由对应的数运算性质推得. 通过引入坐标系, 向量的几何描述可自然地转化为解析描述.

若用公理化法, 我们完全不关心向量及其代数运算的性质, 而是将它们看作未**定义概念** (undefined concept), 对于未定义概念, 我们只知道它们满足一组公理. 我们称这个满足一组预先选定的公理的代数系统为**线性空间** (linear space) 或**线性向量空间** (linear vector space). 我们可在所有的数学分支中找到线性空间的例子, 很

多这样的例子将在第 3 章中讨论. 有向线段代数系统和由分量定义的向量代数系统仅仅是线性空间的两个例子而已.

从公理化的视角来研究向量代数或许在数学上是最令人满意的途径, 这是因为用这种方法的话, 向量的描述完全没有用到坐标系或任何特别的几何表示. 我们也正是用这种方法在第 3 章及以后各章中介绍本书的主要议题 —— 线性代数的基础.

在本章中, 我们主要使用解析法, 同时也会用有向线段来从几何的角度解释许多相关结果, 我们尽量不在证明中使用坐标. 因此本章可用来帮助读者熟悉线性空间的重要实例, 同时也为将在第 3 章中介绍的抽象方法建立基础.

1.2 实 n 元组组成的向量空间

早在 17 世纪, 笛卡儿就开始用数对 (a_1, a_2) 来定位平面上的点, 用三元组 (a_1, a_2, a_3) 来定位空间中的点. 到了 19 世纪, 数学家 A. Carley(1821—1895) 和 H. G. Grassmann(1809—1877) 意识到三元组远不是这种表示法的极限. 我们可以把它推广到四元组 (a_1, a_2, a_3, a_4) 以及更一般的 n 元组

$$(a_1, a_2, \cdots, a_n),$$

其中 $n \geqslant 1$, 这样的 n 元组称为 n **维点**或 n **维向量**, 我们称这 n 个数 $a_1, a_2, \cdots, a_n$ 为该点的**坐标**或该向量的**分量**. 所有的 n 维向量组成的集合称为 n **维空间**或 n 元组组成的向量空间, 用记号 $\mathbb{R}^n$ 表示, 其中 $\mathbb{R}$ 表示该空间中向量的分量均为实数.

研究高维空间并不仅仅是为了推广笛卡儿坐标, 它还提供了自然描述用最多三维坐标无法表示的问题的方法. 例如, 物理中刻画某些事件不能仅用空间的三个坐标, 还得用到时间坐标, 因此, 每一个这样的事件都有一个 $\mathbb{R}^4$ 中的点与之对应. 人们经常碰到有好几个自由度的系统以及牵涉到大量齐次方程的问题, 若使用合适的 n 维空间中的向量并将这些方程用一个向量方程来代替的话, 这些问题将会变得易于分析. 不仅如此, 研究更通用的 n 维空间使我们能将一维、二维和三维等空间的许多性质统一起来, 也就是说我们能获得与维数无关的空间的性质. 这种求一般问题的通用解法的思路也是现代数学的重要精神.

虽然当 $n = 1, 2, 3$ 时, 都有恰当的几何图形促使我们使用一维、二维和三维向量, 并帮助我们描绘这些概念, 但是当 $n > 3$ 时, 就找不到这样的几何图形了, 因此, 高维空间向量代数的研究必须完全通过解析的方法进行.

本章中, 我们用大写字母 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \cdots$ 来表示向量, 用小写字母 $a, b, c, \cdots$ 来表示分量, 例如

$$\boldsymbol{A} = (a_1, a_2, \cdots, a_n).$$

为了将 $\mathbb{R}^n$ 转化为一个代数系统, 我们引入向量的相等、加法和纯量乘法的概

念. 纯量 (scalar) 一词在这里作为实数的同义词使用, 它是 Hamilton 由 stair 或 scale 的拉丁语引入的.

定义 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $\mathbb{R}^n$ 中两个向量, 若它们的各分量分别相等, 则称 A 和 B 相等, 即, 设 $\boldsymbol{A}=(a_1,a_2,\cdots,a_n)$, $\boldsymbol{B}=(b_1,b_2,\cdots,b_n)$, 向量等式 $\boldsymbol{A}=\boldsymbol{B}$ 与如下 n 个纯量等式等价:

$$a_1=b_1,\quad a_2=b_2,\quad \cdots,\quad a_n=b_n.$$

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和 $\boldsymbol{A}+\boldsymbol{B}$ 为将它们的对应分量分别相加所得的向量:

$$\boldsymbol{A}+\boldsymbol{B}=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n).$$

若 c 为一纯量, 则我们定义 $\boldsymbol{A}$ 和 c 的纯量乘法为将 c 与 $\boldsymbol{A}$ 的所有分量相乘所得的向量, 并将它记为 $c\boldsymbol{A}$ 或 $\boldsymbol{A}c$:

$$c\boldsymbol{A}=(ca_1,ca_2,\cdots,ca_n).$$

由上述定义易得如下性质.

定理 1.1 向量加法满足交换律,

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A},$$

也满足结合律,

$$\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}.$$

纯量乘法满足结合律,

$$c(d\boldsymbol{A})=(cd)\boldsymbol{A},$$

并满足如下两条分配律,

$$c(\boldsymbol{A}+\boldsymbol{B})=c\boldsymbol{A}+c\boldsymbol{B},\quad (c+d)\boldsymbol{A}=c\boldsymbol{A}+d\boldsymbol{A}.$$ □

分量全为零的向量称为**零向量**(zero vector), 记为 $\boldsymbol{O}$. 它具有如下性质: 对任意向量 $\boldsymbol{A}$, $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$, 即 $\boldsymbol{O}$ 是向量加法的单位元. 我们将向量 $(-1)\boldsymbol{A}$ 记为 $-\boldsymbol{A}$, 并称它为 $\boldsymbol{A}$ 的**负元**(negative). $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的负元之和称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的**差**, 记作 $\boldsymbol{A}-\boldsymbol{B}$. 因为 $(\boldsymbol{A}+\boldsymbol{B})-\boldsymbol{B}=\boldsymbol{A}$, 所以减法是加法的逆运算. 注意到对任意 $\boldsymbol{A}$, 都有 $0\boldsymbol{A}=\boldsymbol{O}$ 以及 $1\boldsymbol{A}=\boldsymbol{A}$.

细心的读者可能已经注意到二维空间中的向量和复数非常相似, 它们都被定义为有序实数对, 并且它们的加法也完全相同. 因此, 当我们只讨论加法时, 复数和二维向量在代数意义上是相同的, 仅当我们引入乘法时, 它们才会有所区别.

复数的乘法继承了实数乘法的许多性质. 例如, 若两个复数之积为零, 则其中至少有一个为零, 这个性质在用因式分解法求解多项式方程时十分有用. 然而, 人们已经证明了当 $n\geqslant 3$ 时, 无法在同时满足上述性质及交换律、结合律、分配律的前提下在 $\mathbb{R}^n$ 中引入乘法. 但是在 $\mathbb{R}^n$ 中仍然存在满足其中某些性质的几种乘法, 这些乘法将各自在不同的应用中起作用. 例如, 在 1.5 节中, 我们会讨论 $\mathbb{R}^n$ 中两个向量的点积, 这种乘法的结果是一个纯量, 而不是一个向量, 我们会看到两个非零向量的点积可能为零. 另一种称为叉积的乘法将在 2.10 节中讨论, 这种乘法仅当

$n=3$ 时有意义, 它的结果仍然是一个向量, 不过叉积不满足交换律. 此外, 两个非零向量的叉积也可能为零.

1.3 $n \leqslant 3$ 时 n 维向量的几何描述

尽管上述定义与几何完全无关, 不过三维或三维以下空间中的向量及其运算具有非常有趣的几何描述. 我们会在二维空间中用图形来说明这些概念, 并希望读者在三维及一维空间中给出这些概念的几何描述.

设 $\boldsymbol{A}$, $\boldsymbol{B}$ 为空间中的两点, 若我们将其中一点 $\boldsymbol{A}$ 看作起点, 将另一点 $\boldsymbol{B}$ 看作终点, 则我们可以称这对点为一个**几何向量**(geometric vector), 称 $\boldsymbol{A}$ 为**始点**, 称 $\boldsymbol{B}$ 为**终点**. 我们在图像上用从 $\boldsymbol{A}$ 到 $\boldsymbol{B}$ 的有向线段表示这个几何向量, 如图 1.1 所示, 并将这个向量记为 $\overrightarrow{\boldsymbol{AB}}$.

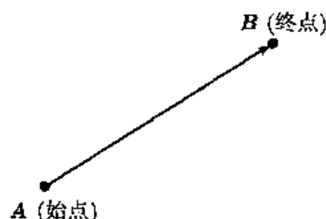


图 1.1 从 $\boldsymbol{A}$ 到 $\boldsymbol{B}$ 的几何向量 $\overrightarrow{\boldsymbol{AB}}$

用几何向量表示诸如力、位移、速度、加速度等既有大小又有方向的物理量非常方便, 有向线段的长度表示该物理量的大小, 而有向线段的指向则表示它的方向. Hamilton 最先从 Carrier 的拉丁语中引入了向量 (vector) 这个词来表示既有大小又有方向的数学量.

假设我们引入以 $\boldsymbol{O}$ 为原点的坐标系, 图 1.2 给出两个几何向量 $\overrightarrow{\boldsymbol{AB}}$ 和 $\overrightarrow{\boldsymbol{CD}}$ 的表示, 其中 $\boldsymbol{B}-\boldsymbol{A}=\boldsymbol{D}-\boldsymbol{C}$. 若用分量表示, 我们有

$$b_1-a_1=d_1-c_1, \quad 且 \quad b_2-a_2=d_2-c_2.$$

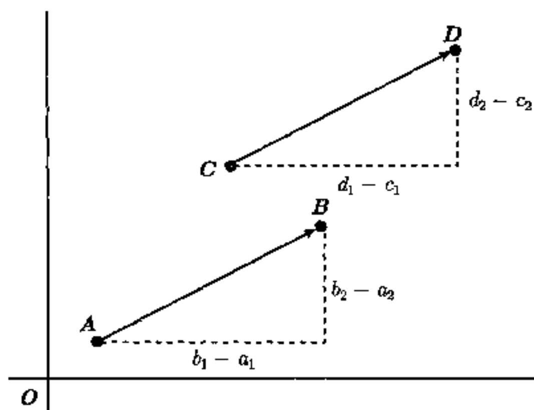


图 1.2 $\overrightarrow{\boldsymbol{AB}}$ 与 $\overrightarrow{\boldsymbol{CD}}$ 等价, 因为 $\boldsymbol{B}-\boldsymbol{A}=\boldsymbol{D}-\boldsymbol{C}$

比较图 1.2 中的两个直角三角形, 我们看到表示 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 的有向线段的长度相等, 而且它们平行且指向相同, 所以我们称这样的两个几何向量**等价**(equivalent), 即当且仅当

$$\boldsymbol{B}-\boldsymbol{A}=\boldsymbol{D}-\boldsymbol{C} \tag{1.1}$$

时, 称 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 等价. 这意味着, 我们在空间中随意平移一个向量所得到的新向量与原向量等价. 此外, 如图 1.3 所示, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}$ 四点确定一个平行四边形.

我们也可将式 (1.1) 写成 $\boldsymbol{A}+\boldsymbol{D}=\boldsymbol{B}+\boldsymbol{C}$ 的形式, 即一个平行四边形的两组**对角顶点** (opposite vertices) 之和相等. 特别当其中的一个顶点 (比如 $\boldsymbol{A}$) 是原点 $\boldsymbol{O}$ 时, 如图 1.4 所示, 那么从 $\boldsymbol{O}$ 到它的对角顶点 $\boldsymbol{D}$ 的几何向量即为向量 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 之和. 我们称这种现象为几何向量的加法遵守**平行四边形法则**(parallelogram law). 考虑到许多诸如力、速度、加速度等物理量之和都遵守平行四边形法则这一重要事实, 向量对物理学十分重要.

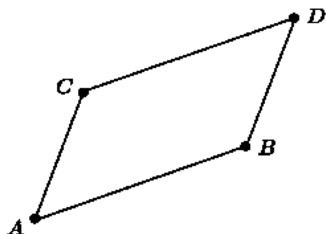


图 1.3 平行四边形两组对角顶点之和相等: $\boldsymbol{A}+\boldsymbol{D}=\boldsymbol{B}+\boldsymbol{C}$

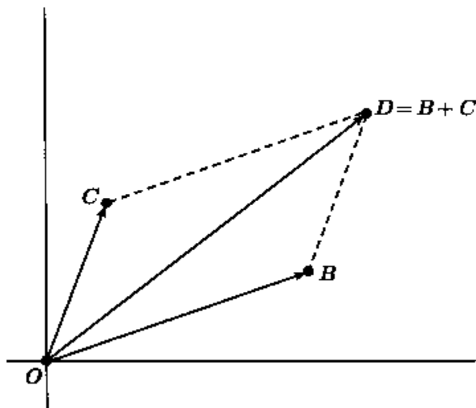


图 1.4 由平行四边形法则给出向量加法的几何解释

为了简便起见, 我们用同一个记号表示 $\mathbb{R}^n(n \leqslant 3)$ 中的点以及从原点到该点的几何向量, 即将 $\overrightarrow{OA}$ 记作 $\boldsymbol{A}$, 将 $\overrightarrow{OB}$ 记作 $\boldsymbol{B}$, 等等. 此外, 我们有时也用 $\boldsymbol{A}$ 表示与 $\overrightarrow{OA}$ 等价的任一几何向量. 例如, 图 1.5 描述了向量减法的几何意义, 其中有两个向量都用 $\boldsymbol{B}-\boldsymbol{A}$ 表示, 注意到这两个向量的大小和方向都相同, 它们是等价的.

图 1.6 描述了向量的纯量乘法的几何意义. 若 $\boldsymbol{B}=c\boldsymbol{A}$, 则几何向量 $\boldsymbol{B}$ 的长度为 c 与 $\boldsymbol{A}$ 的长度之积, 若 c 为正数, 则 $\boldsymbol{B}$ 的方向与 $\boldsymbol{A}$ 的方向相同, 若 c 为负数, 则 $\boldsymbol{B}$ 的方向与 $\boldsymbol{A}$ 的方向相反.

在 $n \leqslant 3$ 时的 n 维空间中向量的几何表示为我们提供了一种定义一般 n 维空间中的向量平行的方法.

定义 令 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 维空间中的向量, 若存在正数 c, 使 $\boldsymbol{B}=c\boldsymbol{A}$, 则称 $\boldsymbol{A}$

和 B 的方向相同; 若存在负数 c, 使 $B = cA$, 则称 A 和 B 的方向相反. 若存在非零数 c, 使 $B = cA$, 则称 A 和 B 平行.

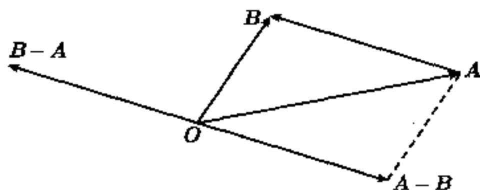


图 1.5　向量减法的几何意义

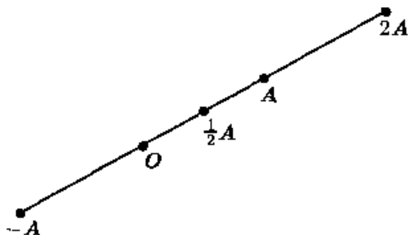


图 1.6　向量的纯量乘法

注意到在这个定义下, 所有的向量都和自己方向相同, 这当然是我们想要的性质. 另外, 此定义也决定了零向量有如下性质: 零向量是唯一一个和自己的负元方向相同的向量, 也是唯一一个和自己方向相反的向量, 它还是唯一一个和零向量平行的向量.

1.4　习　题

1. 令 $A = (1,3,6)$, $B = (4,-3,3)$, $C = (2,1,5)$ 为 $\mathbb{R}^3$ 上三个向量. 求下列各向量的分量.

   (a) $A + B$　(b) $A - B$　(c) $A + B - C$

   (d) $7A - 2B - 3C$　(e) $2A + B - 3C$

2. 画出从原点到点 $A = (2,1)$ 和 $B = (1,3)$ 的两个几何向量的简图, 另外在图上分别画出当 $t = \frac{1}{3}, \frac{1}{2}$, 1, 2, −1, −2 时, 从原点到点 $C = A + tB$ 的几何向量.

3. 若在第 2 题中 $C = tA + B$, 画出从原点到点 C 的几何向量.

4. 令 $A = (2,1)$, $B = (1,3)$, $C = xA + yB$, 其中 x, y 为纯量.

   (a) 当 $x = y = \frac{1}{2}$; $x = \frac{1}{4}, y = \frac{3}{4}$; $x = \frac{1}{3}, y = \frac{2}{3}$; $x = 2, y = -1$; $x = 3, y = -2$; $x = -\frac{1}{2}, y = \frac{3}{2}$; $x = -1, y = 2$ 时, 分别画出从原点到点 C 的几何向量.

   (b) 在 (a) 的基础上, 当 x, y 取遍满足 $x + y = 1$ 的所有实数时, 画出点 C 构成的集合的简图.

   (c) 当 $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ 且 x 和 y 相互独立时, 画出点 C 构成的集合的简图.

   (d) 当 x 取遍区间 $0 \leqslant x \leqslant 1$ 中所有实数, y 取遍全体实数时, 画出点 C 构成的集合的简图.

   (e) 当 x, y 都取遍所有实数时, 画出点 C 构成的集合的简图.

5. 令 $A = (2,1)$, $B = (1,3)$, 证明对 $\mathbb{R}^2$ 中的任意向量 $C = (c_1, c_2)$, 都可找到实数 x 和 y, 使 $C = xA + yB$, 并试用 c_1, c_2 表示 x, y.

6. 令 $A = (1,1,1)$, $B = (0,1,1)$, $C = (1,1,0)$ 为 $\mathbb{R}^3$ 中的三个向量, 并令 $D = xA + yB + zC$, 其中 x, y, z 为纯量.

   (a) 用 x, y, z 表示 D 的各分量.

   (b) 若 $D = O$, 证明 $x = y = z = 0$.

   (c) 求 x, y, z 的值, 使 $D = (1,2,3)$.

7. 令 $A = (1,1,1)$, $B = (0,1,1)$, $C = (2,1,1)$ 为 $\mathbb{R}^3$ 的三个向量, 并令 $D = xA + yB + zC$, 其中 x, y, z 为纯量.

(a) 用 x, y, z 表示 $\boldsymbol{D}$ 的各分量.

(b) 求使 $\boldsymbol{D}=\boldsymbol{O}$ 的不全为零的纯量 x, y, z.

(c) 证明不存在这样的 x, y, z, 使得 $\boldsymbol{D}=(1,2,3)$.

8. 令 $\boldsymbol{A}=(1,1,1,0)$, $\boldsymbol{B}=(0,1,1,1)$, $\boldsymbol{C}=(1,1,0,0)$ 为 $\mathbb{R}^4$ 上的三个向量, 并令 $\boldsymbol{D}=x\boldsymbol{A}+y\boldsymbol{B}+z\boldsymbol{C}$, 其中 x, y, z 为纯量.

   (a) 用 x, y, z 表示 $\boldsymbol{D}$ 的各分量.

   (b) 若 $\boldsymbol{D}=\boldsymbol{O}$, 证明 $x=y=z=0$.

   (c) 求 x, y, z 的值, 使 $\boldsymbol{D}=(1,2,3,4)$.

   (d) 证明不存在这样的 x,y,z, 使得 $\boldsymbol{D}=(1,2,3,4)$.

9. 证明在 $\mathbb{R}^n$ 中平行于同一个向量的两个向量相互平行.

10. 在 $\mathbb{R}^n$ 中给定四个非零向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, 满足 $\boldsymbol{C}=\boldsymbol{A}+\boldsymbol{B}$ 且 $\boldsymbol{A}$ 平行于 $\boldsymbol{D}$. 证明当且仅当 $\boldsymbol{B}$ 平行于 $\boldsymbol{D}$ 时, $\boldsymbol{C}$ 平行于 $\boldsymbol{D}$.

11. (a) 对 $\mathbb{R}^n$ 中的向量证明定理 1.1 中关于向量加法和纯量乘法的所有性质.

    (b) 在平面上用几何向量描述两个分配律 $c(\boldsymbol{A}+\boldsymbol{B})=c\boldsymbol{A}+c\boldsymbol{B}$ 和 $(c+d)\boldsymbol{A}=c\boldsymbol{A}+d\boldsymbol{A}$ 的几何意义.

12. 若 $\mathbb{R}^2$ 中四边形 $OABC$ 是平行四边形, 且 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 是一组对角顶点, 证明 $\boldsymbol{A}+\frac{1}{2}(\boldsymbol{C}-\boldsymbol{A})=\frac{1}{2}\boldsymbol{B}$. 这个式子蕴含的是哪一个几何定理?

1.5 点　　积

现在我们引入 n 维空间中向量的另一种乘法, 这种乘法称为 n 维空间中两个向量的**点积** (dot product) 或**纯量积** (scalar product).

定义　设 $\boldsymbol{A}=(a_1,a_2,\cdots,a_n)$ 与 $\boldsymbol{B}=(b_1,b_2,\cdots,b_n)$ 为 $\mathbb{R}^n$ 中的两个向量, 令

$$\boldsymbol{A}\cdot\boldsymbol{B}=\sum_{k=1}^{n}a_kb_k.$$

$\boldsymbol{A}\cdot\boldsymbol{B}$ 称为向量 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的**点积** (dot product).

由定义可知, 若要计算 $\boldsymbol{A}\cdot\boldsymbol{B}$ 的值, 我们需要将它们对应的分量分别相乘, 然后再将所得之积相加. 向量的这种乘法有如下代数性质.

定理 1.2　对 n 维空间中任意向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 和任意纯量 c, 有

(a) $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{B}\cdot\boldsymbol{A}$　　(交换律),

(b) $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}\cdot\boldsymbol{C}$　　(分配律),

(c) $c(\boldsymbol{A}\cdot\boldsymbol{B})=(c\boldsymbol{A})\cdot\boldsymbol{B}=\boldsymbol{A}\cdot(c\boldsymbol{B})$　　(齐性),

(d) $\boldsymbol{A}\cdot\boldsymbol{A}>0$, 若 $\boldsymbol{A}\neq\boldsymbol{O}$　　(正定性),

(e) $\boldsymbol{A}\cdot\boldsymbol{A}=0$, 若 $\boldsymbol{A}=\boldsymbol{O}$.

证明　由点积的定义易得前三条性质, 我们将它们的证明留给读者作为练习. 现在我们来证明后两条性质. 由定义可知 $\boldsymbol{A}\cdot\boldsymbol{A}=\sum_{k=1}^{n}a_k^2$, 因为等式右边每一项都非负, 所以其和也非负. 更进一步, 当且仅当所有项均为零时, 它们的和为零, 而只有当 $\boldsymbol{A}=\boldsymbol{O}$ 时, 此式的所有项才会都为零. □

在 1.9 节中, 我们还将介绍点积的一个有趣的几何解释. 但是我们在这里先介绍一个关于点积的非常重要的不等式, 这个不等式在向量代数中也是基本的.

定理 1.3(Cauchy-Schwarz 不等式) *若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 维空间中的两个向量, 则*

$$(\boldsymbol{A}\cdot\boldsymbol{B})^2 \leqslant (\boldsymbol{A}\cdot\boldsymbol{A})(\boldsymbol{B}\cdot\boldsymbol{B}) \tag{1.2}$$

更进一步, 当且仅当这两个向量中的一个是另一个的纯量倍数时, 上式中等号成立.

证明 我们用定理 1.1 中的五条性质来证明式 (1.2), 这同时也说明 Cauchy-Schwarz 不等式纯粹只是这五条性质的一个推论, 它不以任何用以推出这五条性质的特定定义为基础.

首先, 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中有一个为零, 则式 (1.2) 显然成立, 所以我们假设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都不为零. 令

$$\boldsymbol{C} = x\boldsymbol{A} - y\boldsymbol{B}, \quad \text{其中 } x = \boldsymbol{B}\cdot\boldsymbol{B},\ y = \boldsymbol{A}\cdot\boldsymbol{B}.$$

由性质 (d) 和 (e) 可得 $\boldsymbol{C}\cdot\boldsymbol{C} \geqslant 0$. 另外, 由性质 (a), (b), (c) 可得

$$\boldsymbol{C}\cdot\boldsymbol{C} = (x\boldsymbol{A} - y\boldsymbol{B})\cdot(x\boldsymbol{A} - y\boldsymbol{B}) = x^2(\boldsymbol{A}\cdot\boldsymbol{A}) - 2xy(\boldsymbol{A}\cdot\boldsymbol{B}) + y^2(\boldsymbol{B}\cdot\boldsymbol{B}).$$

由 $\boldsymbol{C}\cdot\boldsymbol{C} \geqslant 0$ 及 x 和 y 的定义, 我们有

$$(\boldsymbol{B}\cdot\boldsymbol{B})^2(\boldsymbol{A}\cdot\boldsymbol{A}) - 2(\boldsymbol{A}\cdot\boldsymbol{B})^2(\boldsymbol{B}\cdot\boldsymbol{B}) + (\boldsymbol{A}\cdot\boldsymbol{B})^2(\boldsymbol{B}\cdot\boldsymbol{B}) \geqslant 0.$$

又因为 $\boldsymbol{B} \neq \boldsymbol{O}$, 所以由性质 (d) 得 $\boldsymbol{B}\cdot\boldsymbol{B} > 0$, 故我们用 $(\boldsymbol{B}\cdot\boldsymbol{B})$ 同时除上式两边得

$$(\boldsymbol{B}\cdot\boldsymbol{B})(\boldsymbol{A}\cdot\boldsymbol{A}) - (\boldsymbol{A}\cdot\boldsymbol{B})^2 \geqslant 0,$$

即式 (1.2) 成立. 证明过程也表明当且仅当 $\boldsymbol{C} = \boldsymbol{O}$ 时式 (1.2) 中的等号成立. 又因为当且仅当 $x\boldsymbol{A} = y\boldsymbol{B}$ 时, $\boldsymbol{C} = \boldsymbol{O}$, 所以当且仅当 A 和 B 中的一个是另一个的某一纯量倍数时, 式 (1.2) 中的等号成立. □

将 Cauchy-Schwarz 不等式写成分量形式, 我们有

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leqslant \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right).$$

Cauchy-Schwarz 不等式在讨论向量的长度 (范数) 的性质时有非常重要的作用, 我们在下一节中介绍向量的长度 (范数) 的概念.

1.6 向量的模和范数

图 1.7 中画出了平面上从原点到点 $\boldsymbol{A} = (a_1, a_2)$ 的几何向量. 由 Pythagoras 定理, 我们知道 $\boldsymbol{A}$ 的长度由公式

$$\boldsymbol{A} \text{ 的长度} = \sqrt{a_1^2 + a_2^2}$$

给出. 对图 1.8 所示的三维空间中的几何向量 $\boldsymbol{A} = (a_1, a_2, a_3)$, 运用 Pythagoras 定理两次, 我们得到三维空间中几何向量 $\boldsymbol{A}$ 的长度由公式

$$\boldsymbol{A} \text{ 的长度} = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

给出. 注意到在这两种情况下 $\boldsymbol{A}$ 的长度都等于 $(\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}$, 即 $\boldsymbol{A}$ 与自己的点积的平方根. 由此我们可在 n 维空间中定义向量的长度概念.

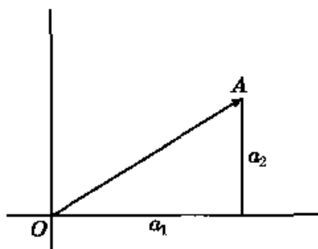


图 1.7 在 $\mathbb{R}^2$ 中, $\boldsymbol{A}$ 的长度为 $\sqrt{a_1^2+a_2^2}$

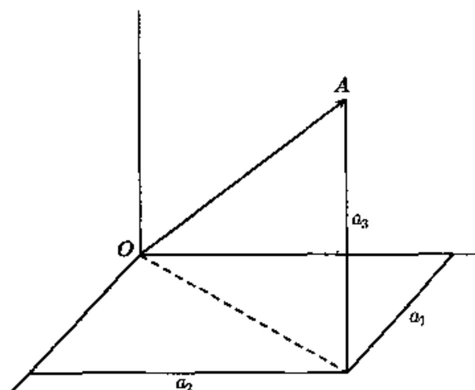


图 1.8 在 $\mathbb{R}^3$ 中, $\boldsymbol{A}$ 的长度为 $\sqrt{a_1^2+a_2^2+a_3^2}$

定义 设 $\boldsymbol{A}$ 为 n 维空间中的向量, 令

$$\|\boldsymbol{A}\|=(\boldsymbol{A}\cdot\boldsymbol{A})^{1/2},$$

$\|\boldsymbol{A}\|$ 称为向量 $\boldsymbol{A}$ 的**长度** (length) 或**范数** (norm).

由点积的基本性质可得范数的对应性质.

定理 1.4 对 n 维空间中任意向量 $\boldsymbol{A}$ 和任意纯量 c, 我们有如下性质:

(a) $\|\boldsymbol{A}\|>0$, 若 $\boldsymbol{A}\neq\boldsymbol{O}$ (正定性),

(b) $\|\boldsymbol{A}\|=0$, 若 $\boldsymbol{A}=\boldsymbol{O}$,

(c) $\|c\boldsymbol{A}\|=|c|\|\boldsymbol{A}\|$ (齐次性).

证明 由定理 1.1 中的 (d) 和 (e) 可直接推得性质 (a) 和 (b). 现在我们来证性质 (c), 由点积的齐性, 我们有

$$\|c\boldsymbol{A}\|=(c\boldsymbol{A}\cdot c\boldsymbol{A})^{1/2}=(c^2\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}=(c^2)^{1/2}(\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}=|c|\|\boldsymbol{A}\|. \qquad \square$$

Cauchy-Schwarz 不等式也可写成范数形式:

$$(\boldsymbol{A}\cdot\boldsymbol{B})^2\leqslant\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2. \tag{1.3}$$

对上式两边同时取正平方根, 我们也可将 Cauchy-Schwarz 不等式写成如下等价形式:

$$|\boldsymbol{A}\cdot\boldsymbol{B}|\leqslant\|\boldsymbol{A}\|\|\boldsymbol{B}\|. \tag{1.4}$$

现在我们用 Cauchy-Schwarz 不等式推导三角不等式.

定理 1.5 (三角不等式) 对 n 维空间中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 我们有

$$\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|.$$

更进一步, 当且仅当 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$ 或存在纯量 $c>0$, 使得 $\boldsymbol{B}=c\boldsymbol{A}$ 时, 上式中等号成立.

注 图 1.9 给出了三角不等式的几何意义, 它表示一个三角形的一条边长不会比另两条边长之和大.

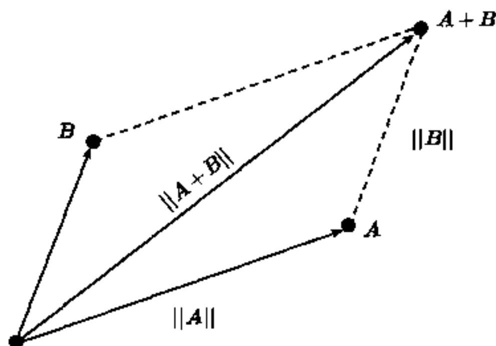


图 1.9 三角不等式 $\|\boldsymbol{A}+\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\|+\|\boldsymbol{B}\|$ 的几何意义

证明 我们将三角不等式写成如下等价形式:

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2 \leqslant (\|\boldsymbol{A}\|+\|\boldsymbol{B}\|)^2. \tag{1.5}$$

上式左边可写成

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2=(\boldsymbol{A}+\boldsymbol{B})\cdot(\boldsymbol{A}+\boldsymbol{B})=\boldsymbol{A}\cdot\boldsymbol{A}+2\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{B}\cdot\boldsymbol{B}=\|\boldsymbol{A}\|^2+2\boldsymbol{A}\cdot\boldsymbol{B}+\|\boldsymbol{B}\|^2,$$

而右边可写成

$$(\|\boldsymbol{A}\|+\|\boldsymbol{B}\|)^2=\|\boldsymbol{A}\|^2+2\|\boldsymbol{A}\|\|\boldsymbol{B}\|+\|\boldsymbol{B}\|^2.$$

比较以上两式, 我们知道当且仅当

$$\boldsymbol{A}\cdot\boldsymbol{B} \leqslant \|\boldsymbol{A}\|\|\boldsymbol{B}\| \tag{1.6}$$

成立时, 式 (1.5) 成立. 又因为 $\boldsymbol{A}\cdot\boldsymbol{B} \leqslant |\boldsymbol{A}\cdot\boldsymbol{B}|$, 故由式 (1.4) 知式 (1.6) 成立. 这说明可由 Cauchy-Schwarz 不等式推得三角不等式.

此外, 也可由三角不等式推得 Cauchy-Schwarz 不等式, 即若三角不等式成立, 则对 $\boldsymbol{A}$ 和 $-\boldsymbol{A}$, 式 (1.6) 也成立, 由此可知式 (1.3) 成立.

另外, 若式 (1.5) 中等号成立, 则 $\boldsymbol{A}\cdot\boldsymbol{B}=\|\boldsymbol{A}\|\|\boldsymbol{B}\|$, 所以存在纯量 c, 使得 $\boldsymbol{B}=c\boldsymbol{A}$. 故 $\boldsymbol{A}\cdot\boldsymbol{B}=c\|\boldsymbol{A}\|^2$, $\|\boldsymbol{A}\|\|\boldsymbol{B}\|=|c|\|\boldsymbol{A}\|^2$. 若 $\boldsymbol{A}\neq\boldsymbol{O}$, 则有 $c=|c|$, 所以 $c\geqslant 0$. 又若 $\boldsymbol{B}\neq\boldsymbol{O}$, 有 $\boldsymbol{B}=c\boldsymbol{A}$ 且 $c>0$. □

1.7 向量的正交

在证明三角不等式 (定理 1.5) 的过程中, 我们得到公式

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2+2\boldsymbol{A}\cdot\boldsymbol{B}, \tag{1.7}$$

这个公式对 n 维空间中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都成立. 图 1.10 画出了空间中相互垂直的两个向量, 它们给出了一个直角边长分别为 $\|\boldsymbol{A}\|$ 和 $\|\boldsymbol{B}\|$, 斜边长为 $\|\boldsymbol{A}+\boldsymbol{B}\|$ 的直角三角形.

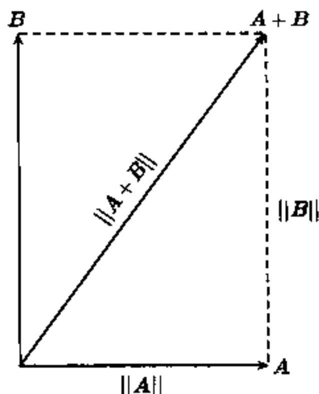


图 1.10 满足 Pythagoras 恒等式 $\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2$ 的两个相互垂直的向量

由于 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相互垂直, Pythagoras 定理告诉我们

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2.$$

比较上式和式 (1.7), 有 $\boldsymbol{A}\cdot\boldsymbol{B}=0$, 即在 $\mathbb{R}^2$ 上, 两个相互垂直的向量的点积为零. 我们可对 n 维空间中的向量给出垂直的定义.

定义 对 $\mathbb{R}^n$ 上两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 当 $\boldsymbol{A}\cdot\boldsymbol{B}=0$ 时, 称它们**互相垂直** (perpendicular) 或**正交** (orthogonal).

等式 (1.7) 告诉我们, 对 $\mathbb{R}^n$ 上两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 当且仅当 $\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2$ 时, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 正交. 我们称此式为 $\mathbb{R}^n$ 上的 Pythagoras 恒等式.

1.8 习题

1. 令 $\boldsymbol{A}=(1,2,3,4)$, $\boldsymbol{B}=(-1,2,-3,0)$, $\boldsymbol{C}=(0,1,0,1)$ 为 $\mathbb{R}^4$ 中三个向量, 计算下列各点积的值:
   (a) $\boldsymbol{A}\cdot\boldsymbol{B}$ (b) $\boldsymbol{B}\cdot\boldsymbol{C}$
   (c) $\boldsymbol{A}\cdot\boldsymbol{C}$ (d) $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})$
   (e) $(\boldsymbol{A}-\boldsymbol{B})\cdot\boldsymbol{C}$.
2. 给定向量 $\boldsymbol{A}=(2,4,-7)$, $\boldsymbol{B}=(2,6,3)$, $\boldsymbol{C}=(3,4,-5)$. 下列各式都只有一种添加括号的方式使它们成为有意义的表达式. 试为下列各式添加括号并计算它们的值:
   (a) $\boldsymbol{A}\cdot\boldsymbol{B}\boldsymbol{C}$ (b) $\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{C}$
   (c) $\boldsymbol{A}+\boldsymbol{B}\cdot\boldsymbol{C}$ (d) $\boldsymbol{A}\boldsymbol{B}\cdot\boldsymbol{C}$
   (e) $\boldsymbol{A}/\boldsymbol{B}\cdot\boldsymbol{C}$.
3. 判断关于 $\mathbb{R}^n$ 中向量的下述命题是否为真并给出证明: 若 $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{A}\cdot\boldsymbol{C}$, 且 $\boldsymbol{A}\neq\boldsymbol{O}$, 则 $\boldsymbol{B}=\boldsymbol{C}$.
4. 判断关于 $\mathbb{R}^n$ 中向量的下述命题是否为真并给出证明: 若对任意 $\boldsymbol{B}$, 有 $\boldsymbol{A}\cdot\boldsymbol{B}=0$, 则 $\boldsymbol{A}=\boldsymbol{O}$.
5. 设 $\boldsymbol{A}=(2,1,-1)$, $\boldsymbol{B}=(1,-1,2)$, 求 $\mathbb{R}^3$ 中的非零向量 $\boldsymbol{C}$, 使得 $\boldsymbol{A}\cdot\boldsymbol{C}=\boldsymbol{B}\cdot\boldsymbol{C}=0$.
6. 设 $\boldsymbol{A}=(1,-2,3)$, $\boldsymbol{B}=(3,1,2)$, 求纯量 x 和 y, 使得 $\boldsymbol{C}=x\boldsymbol{A}+y\boldsymbol{B}$ 为非零向量, 且 $\boldsymbol{C}\cdot\boldsymbol{B}=0$.
7. 令 $\boldsymbol{A}=(2,-1,2)$, $\boldsymbol{B}=(1,2,-2)$, 求 $\mathbb{R}^3$ 中同时满足以下条件的向量 $\boldsymbol{C}$ 和 $\boldsymbol{D}$: $\boldsymbol{A}=\boldsymbol{C}+\boldsymbol{D}$, $\boldsymbol{B}\cdot\boldsymbol{D}=0$, 且 $\boldsymbol{C}$ 平行于 $\boldsymbol{B}$.
8. 令 $\boldsymbol{A}=(1,2,3,4,5)$, $\boldsymbol{B}=(1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{5})$, 求 $\mathbb{R}^5$ 中同时满足以下条件的向量 $\boldsymbol{C}$ 和 $\boldsymbol{D}$: $\boldsymbol{B}=\boldsymbol{C}+2\boldsymbol{D}$, $\boldsymbol{D}\cdot\boldsymbol{A}=0$, 且 $\boldsymbol{C}$ 平行于 $\boldsymbol{A}$.
9. 令 $\boldsymbol{A}=(2,-1,5)$, $\boldsymbol{B}=(-1,-2,3)$, $\boldsymbol{C}=(1,-1,1)$ 为 $\mathbb{R}^3$ 中三个向量, 计算下列各向量的范数:
   (a) $\boldsymbol{A}+\boldsymbol{B}$ (b) $\boldsymbol{A}-\boldsymbol{B}$
   (c) $\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{C}$ (d) $\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{C}$.
10. 对以下各向量 $\boldsymbol{A}$, 求 $\mathbb{R}^2$ 中的向量 $\boldsymbol{B}$, 使得 $\boldsymbol{B}\cdot\boldsymbol{A}=0$ 且 $\|\boldsymbol{A}\|=\|\boldsymbol{B}\|$:
    (a) $\boldsymbol{A}=(1,1)$ (b) $\boldsymbol{A}=(1,-1)$
    (c) $\boldsymbol{A}=(2,-3)$ (d) $\boldsymbol{A}=(a,b)$.
11. 令 $\boldsymbol{A}=(1,-2,3)$, $\boldsymbol{B}=(3,1,2)$ 为 $\mathbb{R}^3$ 中的两个向量, 分别求平行于以下各向量且长度为 1 的向量 $\boldsymbol{C}$:

(a) $\boldsymbol{A}+\boldsymbol{B}$　　(b) $\boldsymbol{A}-\boldsymbol{B}$
(c) $\boldsymbol{A}+2\boldsymbol{B}$　　(d) $\boldsymbol{A}-2\boldsymbol{B}$
(e) $2\boldsymbol{A}-\boldsymbol{B}$.

12. 令 $\boldsymbol{A}=(4,1,-3)$, $\boldsymbol{B}=(1,2,2)$, $\boldsymbol{C}=(1,2,-2)$, $\boldsymbol{D}=(2,1,2)$, $\boldsymbol{E}=(2,-2,-1)$ 为 $\mathbb{R}^3$ 中向量, 找出其中所有相互正交的向量对.
13. 对于下列向量 $\boldsymbol{A}$, 求 $\mathbb{R}^2$ 中所有与 $\boldsymbol{A}$ 正交且长度与 $\boldsymbol{A}$ 相同的向量:
(a) $\boldsymbol{A}=(1,2)$　　(b) $\boldsymbol{A}=(1,-2)$
(c) $\boldsymbol{A}=(2,-1)$　　(d) $\boldsymbol{A}=(-2,1)$.
14. 设 $\boldsymbol{A}=(2,-1,1)$, $\boldsymbol{B}=(3,-4,-4)$, 求三维空间中的点 $\boldsymbol{C}$, 使得 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 构成一个直角三角形.
15. 设 $\boldsymbol{A}=(1,-1,2)$, $\boldsymbol{B}=(2,1,-1)$, 求同时与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 正交的 $\mathbb{R}^3$ 中的非零向量 $\boldsymbol{C}$.
16. 令 $\boldsymbol{A}=(1,2)$, $\boldsymbol{B}=(3,4)$ 为 $\mathbb{R}^2$ 中的两个向量, 求 $\mathbb{R}^2$ 中两向量 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 满足 $\boldsymbol{A}=\boldsymbol{P}+\boldsymbol{Q}$, $\boldsymbol{P}$ 平行于 $\boldsymbol{B}$, $\boldsymbol{Q}$ 与 $\boldsymbol{B}$ 正交.
17. 令 $\boldsymbol{A}=(1,2,3,4)$, $\boldsymbol{B}=(1,1,1,1)$ 为 $\mathbb{R}^4$ 中的两个向量, 求 $\mathbb{R}^4$ 中两向量 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 满足 $\boldsymbol{A}=\boldsymbol{P}+\boldsymbol{Q}$, $\boldsymbol{P}$ 平行于 $\boldsymbol{B}$, $\boldsymbol{Q}$ 与 $\boldsymbol{B}$ 正交.
18. 给定 $\mathbb{R}^3$ 中向量 $\boldsymbol{A}=(2,-1,1)$, $\boldsymbol{B}=(1,2,-1)$, $\boldsymbol{C}=(1,1,-2)$, 求所有长度为 1, 形如 $x\boldsymbol{B}+y\boldsymbol{C}$ 且与 $\boldsymbol{A}$ 正交的向量 $\boldsymbol{D}$.
19. 证明: 对 $\mathbb{R}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 我们有恒等式
$$\|\boldsymbol{A}+\boldsymbol{B}\|^2-\|\boldsymbol{A}-\boldsymbol{B}\|^2=4\boldsymbol{A}\cdot\boldsymbol{B}.$$
这个恒等式告诉我们, 当且仅当 $\|\boldsymbol{A}+\boldsymbol{B}\|=\|\boldsymbol{A}-\boldsymbol{B}\|$ 时, $\boldsymbol{A}\cdot\boldsymbol{B}=0$. 从几何意义上说, 若考虑 $\mathbb{R}^2$ 中的向量, 这个恒等式说明当且仅当一个平行四边形为矩形时, 它的两条对角线的长度相等.
20. 证明: 对 $\mathbb{R}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 我们有
$$\|\boldsymbol{A}+\boldsymbol{B}\|^2+\|\boldsymbol{A}-\boldsymbol{B}\|^2=2\|\boldsymbol{A}\|^2+2\|\boldsymbol{B}\|^2.$$
由这个恒等式可以推出平行四边形的边和对角线的哪一个几何定理?
21. 由下述几何定理可推出关于三个向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的一个恒等式, 试猜测这个恒等式并证明它对 $\mathbb{R}^n$ 中的所有向量都成立.
定理　*任意四边形的边长的平方和与其对角线长的平方和之差等于其对角线中点连线长的平方的 4 倍.*
22. 设 $\mathbb{R}^n$ 中向量 $\boldsymbol{A}$ 的长度为 6, $\mathbb{R}^n$ 中另一个向量 $\boldsymbol{B}$ 有性质: 对任意两个纯量 x 和 y, 向量 $x\boldsymbol{A}+y\boldsymbol{B}$ 与向量 $4y\boldsymbol{A}-9x\boldsymbol{B}$ 正交, 求 $\boldsymbol{B}$ 和 $2\boldsymbol{A}+3\boldsymbol{B}$ 的长度.
23. 给定 $\mathbb{R}^5$ 中两个向量 $\boldsymbol{A}=(1,2,3,4,5)$, $\boldsymbol{B}=(1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{5})$ 求 $\mathbb{R}^5$ 中的向量 $\boldsymbol{C}$ 和 $\boldsymbol{D}$, 使得 $\boldsymbol{C}$ 平行于 $\boldsymbol{A}$, $\boldsymbol{D}$ 与 $\boldsymbol{A}$ 正交, 且 $\boldsymbol{C}+\boldsymbol{D}=\boldsymbol{B}$.
24. 给定 $\mathbb{R}^n$ 中两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 其中 $\boldsymbol{A}\neq\boldsymbol{O}$, 证明 $\mathbb{R}^n$ 中存在向量 $\boldsymbol{C}$ 和 $\boldsymbol{D}$, 使得 $\boldsymbol{C}$ 平行于 $\boldsymbol{A}$, $\boldsymbol{D}$ 与 $\boldsymbol{A}$ 正交, 且 $\boldsymbol{C}+\boldsymbol{D}=\boldsymbol{B}$. 并用 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 表示 $\boldsymbol{C}$ 和 $\boldsymbol{D}$.
25. 判断下列关于 $\mathbb{R}^n$ 中向量的命题是否为真, 并给出证明:
(a) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交, 则对所有实数 x, $\|\boldsymbol{A}+x\boldsymbol{B}\|\geqslant\|\boldsymbol{A}\|$.
(b) 若对所有实数 x 有 $\|\boldsymbol{A}+x\boldsymbol{B}\|\geqslant\|\boldsymbol{A}\|$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交.

1.9 投影 · n 维空间中向量的夹角

平面上两个向量的点积有一个非常有趣的几何解释. 图 1.11a 画出了两个非零向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 之间的夹角为 θ 的情况. 在这个例子中, 我们有 $0\leqslant\theta<\frac{1}{2}\pi$, 图 1.11b 画出了同一个向量 $\boldsymbol{A}$ 以及两个和为 $\boldsymbol{A}$ 且相互垂直的向量. 这两个向量中的一个

为 $t\boldsymbol{B}$, 它是 $\boldsymbol{B}$ 的一个纯量倍数, 我们称它为$\boldsymbol{A}$ 在 $\boldsymbol{B}$ 上的投影. 在本例中, 因为 $0<\theta\leqslant\frac{1}{2}\pi$, 所以 t 为正数.

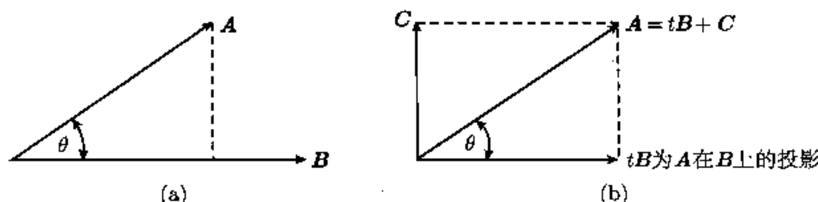


图 1.11 向量 $t\boldsymbol{B}$ 是 $\boldsymbol{A}$ 在 $\boldsymbol{B}$ 上的投影

我们可以用 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的点积来表示 t. 首先, 我们知道 $t\boldsymbol{B}+\boldsymbol{C}=\boldsymbol{A}$, 等式两边同时点乘 $\boldsymbol{B}$, 我们有

$$t\boldsymbol{B}\cdot\boldsymbol{B}+\boldsymbol{C}\cdot\boldsymbol{B}=\boldsymbol{A}\cdot\boldsymbol{B}.$$

由于 $\boldsymbol{C}$ 与 $\boldsymbol{B}$ 垂直, 故有 $\boldsymbol{C}\cdot\boldsymbol{B}=0$, 所以 $t\boldsymbol{B}\cdot\boldsymbol{B}=\boldsymbol{A}\cdot\boldsymbol{B}$, 即我们有

$$t=\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\boldsymbol{B}\cdot\boldsymbol{B}}=\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\|\boldsymbol{B}\|^2}. \tag{1.8}$$

另一方面, 纯量 t 和 θ 有一个简单关系. 事实上, 观察图 1.11b, 我们有

$$\cos\theta=\frac{\|t\boldsymbol{B}\|}{\|\boldsymbol{A}\|}=\frac{t\|\boldsymbol{B}\|}{\|\boldsymbol{A}\|}.$$

将式 (1.8) 代入上式, 我们有

$$\cos\theta=\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|}. \tag{1.9}$$

上式也可写成

$$\boldsymbol{A}\cdot\boldsymbol{B}=\|\boldsymbol{A}\|\|\boldsymbol{B}\|\cos\theta.$$

即 $\mathbb{R}^2$ 中两个非零向量的点积等于它们的长度以及它们夹角的余弦之积.

等式 (1.9) 为我们提供了定义 n 维空间中角的一种方法. 对 $\mathbb{R}^n$ 上任意两个非零向量, 由 Cauchy-Schwarz 不等式 (式 (1.4)), 我们知道式 (1.9) 右边的绝对值 $\leqslant 1$, 即

$$-1\leqslant\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|}\leqslant 1.$$

所以有且仅有一个区间 $[0,\pi]$ 上的实数 θ 使式 (1.9) 成立, 我们将这个 θ 定义为 A 和 B 的夹角. 本节的所有讨论可总结为如下定义.

定义 令 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $\mathbb{R}^n$ 上两个向量, 且 $\boldsymbol{B}\neq\boldsymbol{O}$, 则向量 $t\boldsymbol{B}$ 称为 $\boldsymbol{A}$ 在 $\boldsymbol{B}$ 上的投影, 其中

$$t=\frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\boldsymbol{B}\cdot\boldsymbol{B}}.$$

若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 均不为零, 则 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的夹角 θ 定义如下:

$$\theta = \arccos \frac{\boldsymbol{A}\cdot\boldsymbol{B}}{\boldsymbol{B}\cdot\boldsymbol{B}}.$$

注 由反余弦的定义可知 θ 在区间 $[0,\pi]$ 上.

1.10 单位坐标向量

在第 0 章中, 我们知道任意复数 (a,b) 都可表示为 $a+b\mathrm{i}$ 的形式, 其中 i 表示复数 $(0,1)$. 类似地, $\mathbb{R}^2$ 上任意向量 (a,b) 都可表示为如下形式:

$$(a,b)=a(1,0)+b(0,1).$$

上式中分别与 a 和 b 相乘的向量 $(1,0)$ 和 $(0,1)$ 称为**单位坐标向量**. 现在我们引入 $\mathbb{R}^n$ 中单位坐标向量的概念.

定义 $\mathbb{R}^n$ 上的 n 维向量 $\boldsymbol{I}_1=(1,0,\cdots,0)$, $\boldsymbol{I}_2=(0,1,0,\cdots,0)$, $\cdots$, $\boldsymbol{I}_n=(0,0,\cdots,1)$ 称为单位坐标向量, 其中 $\boldsymbol{I}_k$ 的第 k 个分量为 1, 其他所有分量均为 0.

因为所有这样的向量的长度都为 1, 所以称它们为**单位向量**. 注意到这些向量两两正交, 即任意两个不同的单位坐标向量的点积都为零:

$$\boldsymbol{I}_k\cdot\boldsymbol{I}_j=0,\quad 若\quad k\neq j.$$

定理 1.6 $\mathbb{R}^n$ 上任意向量 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 都可用如下形式表示:

$$\boldsymbol{X}=x_1\boldsymbol{I}_1+\cdots+x_n\boldsymbol{I}_n=\sum_{k=1}^{n}x_k\boldsymbol{I}_k.$$

而且这种表示法唯一, 即若有

$$\boldsymbol{X}=\sum_{k=1}^{n}x_k\boldsymbol{I}_k\quad 且\quad \boldsymbol{X}=\sum_{k=1}^{n}y_k\boldsymbol{I}_k,$$

则对所有 $k=1,2,\cdots,n$, 都有 $x_k=y_k$.

证明 由向量的加法和纯量乘法的定义易得第一个结论, 而由向量相等的定义可得第二个结论. □

形如 $\sum\limits_{k=1}^{n}c_k\boldsymbol{A}_k$ 的向量称为向量 $\boldsymbol{A}_1$, $\cdots$, $\boldsymbol{A}_n$ 的一个**线性组合**. 定理 1.6 告诉我们 $\mathbb{R}^n$ 中的所有向量都可表示为单位坐标向量的线性组合. 我们称这 n 个单位坐标向量**生成** (span) 空间 $\mathbb{R}^n$. 由于任意向量若用 $\boldsymbol{I}_1$, $\cdots$, $\boldsymbol{I}_n$ 的线性组合来表示, 其表示法唯一, 我们也称这样的生成唯一. 其他一些向量组也可唯一地生成 $\mathbb{R}^n$, 在 1.12 节中, 我们将会讨论这样的向量集合.

我们通常用记号 $\boldsymbol{i}$ 和 $\boldsymbol{j}$(i, j 的粗斜体) 表示二维空间中的单位坐标向量 $\boldsymbol{I}_1$ 和 $\boldsymbol{I}_2$; 在三维空间中, 用记号 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 表示 $\boldsymbol{I}_1$, $\boldsymbol{I}_2$, $\boldsymbol{I}_3$. 有时人们也会在这些记号上加一条短线或一个箭头来表示相同的概念. 当 $n=3$ 时, 定理 1.6 的几何意义如图 1.12 所示.

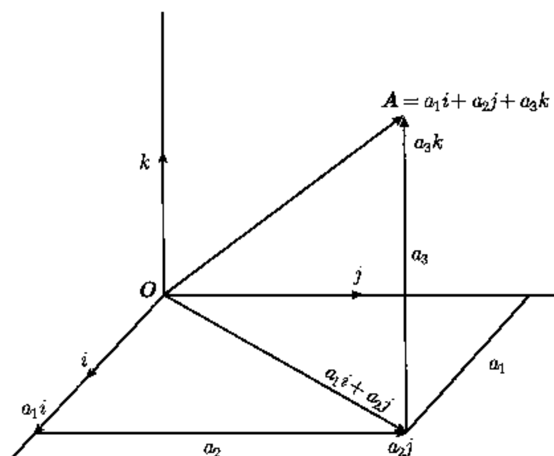


图 1.12 用 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 的线性组合表示 $\mathbb{R}^3$ 中的向量 $\boldsymbol{A}$

当用单位坐标向量的线性组合表示参与代数运算的向量时, 运算结果也可表示为单位坐标向量的线性组合, 且它的各分量为对参与运算的各向量的对应分量进行相同的代数运算所得的结果. 进行这样的计算时, 任何一步结果的各分量都可通过计算各单位坐标向量的系数得到. 例如, 要计算两个向量 $\boldsymbol{A} = (a_1, \cdots, a_n)$, $\boldsymbol{B} = (b_1, \cdots, b_n)$ 之和, 首先将它们写成

$$\boldsymbol{A} = \sum_{k=1}^{n} a_k \boldsymbol{I}_k, \quad \boldsymbol{B} = \sum_{k=1}^{n} b_k \boldsymbol{I}_k,$$

再由有限项之和的线性性可得

$$\boldsymbol{A} + \boldsymbol{B} = \sum_{k=1}^{n} a_k \boldsymbol{I}_k + \sum_{k=1}^{n} b_k \boldsymbol{I}_k = \sum_{k=1}^{n} (a_k + b_k) \boldsymbol{I}_k.$$

上式等号右边的 $\boldsymbol{I}_k$ 的系数即为 $\boldsymbol{A} + \boldsymbol{B}$ 的第 k 个分量.

1.11 习 题

1. 设 $\boldsymbol{A} = (1, 2, 3)$, $\boldsymbol{B} = (1, 2, 2)$, 求 $\boldsymbol{A}$ 在 $\boldsymbol{B}$ 上的投影.
2. 设 $\boldsymbol{A} = (4, 3, 2, 1)$, $\boldsymbol{B} = (1, 1, 1, 1)$, 求 $\boldsymbol{A}$ 在 $\boldsymbol{B}$ 上的投影.
3. (a) 令 $\boldsymbol{A} = (6, 3, -2)$, 并令 a, b, c 分别表示 A 与单位坐标向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 的夹角, 求 $\cos a, \cos b, \cos c$, 它们称为 $\boldsymbol{A}$ 的**方向余弦** (direction cosine).

   (b) 求 $\mathbb{R}^3$ 中平行于 $\boldsymbol{A}$ 且长度为 1 的所有向量.
4. 证明向量 $\boldsymbol{A} = (1, 2, 1)$ 和 $\boldsymbol{B} = (2, 1, -1)$ 的夹角的大小为 $\boldsymbol{C} = (1, 4, 1)$ 和 $\boldsymbol{D} = (2, 5, 5)$ 的夹角的大小的两倍.
5. 给定三维空间中三点 $(2, -1, 1)$, $(1, -3, -5)$, $(3, -4, -4)$, 用向量方法求以这三个点为顶点的三角形的三个角的余弦.

6. $\mathbb{R}^3$ 中三个向量 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 满足以下性质: $\|\boldsymbol{A}\| = \|\boldsymbol{C}\| = 5$, $\|\boldsymbol{B}\| = 1$, $\|\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{C}\| = \|\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C}\|$. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的夹角为 $\pi/8$, 求 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的夹角.
7. 给定 $\mathbb{R}^n$ 中三个非零向量 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, 设 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 的夹角和 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 的夹角相等, 证明 $\boldsymbol{B}$ 与向量 $\|\boldsymbol{B}\|\boldsymbol{A} - \|\boldsymbol{A}\|\boldsymbol{B}$ 正交.
8. 令 θ 为 $\mathbb{R}^n$ 中两个向量 $\boldsymbol{A} = (1,1,\cdots,1)$ 和 $\boldsymbol{B} = (1,2,\cdots,n)$ 的夹角, 求当 $n\to\infty$ 时 θ 的极限值.
9. 在习题 8 中, 若 $\boldsymbol{A} = (2,4,6,\cdots,2n)$, $\boldsymbol{B} = (1,3,5,\cdots,2n-1)$, 求当 $n\to\infty$ 时 θ 的极限值.
10. 给定 $\mathbb{R}^2$ 中向量 $\boldsymbol{A} = (\cos\theta, -\sin\theta)$, $\boldsymbol{B} = (\sin\theta, \cos\theta)$,
    (a) 证明 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是正交的单位向量. 画出 $\theta = \pi/6$ 时 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的简图.
    (b) 求 $\mathbb{R}^2$ 中使得 $(x,y) = x\boldsymbol{A} + y\boldsymbol{B}$ 的所有向量 (x,y). 注意要考虑 θ 的所有情况.
11. 用向量法证明菱形的对角线相互垂直.
12. 用向量 $(\cos a, \sin a)$ 和 $(\cos b, \sin b)$ 的点积推导三角恒等式 $\cos(a-b) = \cos a\cos b + \sin a \sin b$.
13. 设 θ 为 $\mathbb{R}^n$ 中两个非零向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的夹角, 证明
    $$\|\boldsymbol{A}-\boldsymbol{B}\|^2 = \|\boldsymbol{A}\|^2 + \|\boldsymbol{B}\|^2 - 2\|\boldsymbol{A}\|\|\boldsymbol{B}\|\cos\theta.$$
    考虑在 $\mathbb{R}^2$ 中的几何意义时, 上式即为三角学中的余弦定理.
14. 对向量 $\boldsymbol{A} = (a_1,\cdots,a_n)$ 和 $\boldsymbol{B} = (b_1,\cdots,b_n)$, 若我们用公式
    $$\boldsymbol{A}\cdot\boldsymbol{B} = \sum_{k=1}^{n} |a_k b_k|$$
    代替公式 $\boldsymbol{A}\cdot\boldsymbol{B} = \sum\limits_{k=1}^{n} a_k b_k$ 来定义向量的点积, 那么定理 1.2 中的哪几条性质仍然成立? Cauchy-Schwarz 不等式是否仍然成立?
15. 对 $\mathbb{R}^2$ 中向量 $\boldsymbol{A} = (a_1,a_2)$ 和 $\boldsymbol{B} = (b_1,b_2)$, 若我们用公式
    $$\boldsymbol{A}\cdot\boldsymbol{B} = 2a_1b_1 + a_2b_2 + a_1b_2 + a_2b_1$$
    定义向量的点积, 证明定理 1.2 中的所有性质仍然成立. 此外 Cauchy-Schwarz 不等式是否仍然成立?
16. 对 $\mathbb{R}^3$ 中向量 $\boldsymbol{A} = (a_1,a_2,a_3)$ 和 $\boldsymbol{B} = (b_1,b_2,b_3)$, 若我们用公式
    $$\boldsymbol{A}\cdot\boldsymbol{B} = 2a_1b_1 + a_2b_2 + a_3b_3 + a_1b_3 + a_3b_1$$
    定义向量的点积, 证明定理 1.2 中的所有性质仍然成立. 此外 Cauchy-Schwarz 不等式是否仍然成立?
17. 对向量 $\boldsymbol{A} = (a_1,\cdots,a_n)$, 若我们用公式
    $$\|\boldsymbol{A}\| = \sum_{k=1}^{n} |a_k|$$
    代替公式 $\|\boldsymbol{A}\| = (\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}$ 定义向量的范数.
    (a) 证明这样定义的范数满足定理 1.4 和定理 1.5 中的所有性质.
    (b) 在 $\mathbb{R}^2$ 中考虑这样的定义, 并画出所有范数为 1 的点 (x,y) 的集合的简图.
    (c) 证明: 若用公式
    $$\|\boldsymbol{A}\| = \left|\sum_{k=1}^{n} a_k\right|$$
    定义向量的范数, 定理 1.4 和定理 1.5 中的所有性质仍然成立.
18. 对向量 $\boldsymbol{A} = (a_1,\cdots,a_n)$, 若我们用公式 $\|\boldsymbol{A}\| = \max_{1\leqslant k\leqslant n}|a_k|$ 定义向量的范数, 此公式的右边表示 n 个数 $|a_1|, |a_2|, \cdots, |a_n|$ 的最大值.
    (a) 在这样的定义下, 定理 1.4 和定理 1.5 中的哪些性质仍然成立?
    (b) 在这样的定义下, 画出 $\mathbb{R}^2$ 中范数为 1 的所有点 (x,y) 的集合的简图.
19. 设 $\mathbb{R}^n$ 中向量 $\boldsymbol{A} = (a_1,\cdots,a_n)$, 由下式可定义两种范数:
    $$\|\boldsymbol{A}\|_1 = \sum_{k=1}^{n} |a_k|, \quad \|\boldsymbol{A}\|_2 = \max_{1\leqslant k\leqslant n} |a_k|.$$

证明 $\|A\|_2 \leqslant \|A\| \leqslant \|A\|_1$, 并在平面上给出此不等式的几何意义.

20. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 n 维空间中的两个点, 我们称 $\|\boldsymbol{A}-\boldsymbol{B}\|$ 为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 之间的距离, 记作 $d(\boldsymbol{A},\boldsymbol{B})$, 证明这样定义的距离有如下性质.
   (a) $d(\boldsymbol{A},\boldsymbol{B})=d(\boldsymbol{B},\boldsymbol{A})$.
   (b) 当且仅当 $\boldsymbol{A}=\boldsymbol{B}$ 时, $d(\boldsymbol{A},\boldsymbol{B})=0$.
   (c) $d(\boldsymbol{A},\boldsymbol{B})\leqslant d(\boldsymbol{A},\boldsymbol{C})+d(\boldsymbol{C},\boldsymbol{B})$.

1.12 有限向量组的线性生成集

令 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 为包含 k 个 $\mathbb{R}^n$ 中向量的非空集合, 其中向量的个数 k 可能小于、等于或大于空间的维数 n. 若 $\mathbb{R}^n$ 中的向量 $\boldsymbol{X}$ 可表示为 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k$ 的一个线性组合, 设

$$\boldsymbol{X}=\sum_{i=1}^{k}c_i\boldsymbol{A}_i,$$

则称集合 S **生成**向量 $\boldsymbol{X}$.

定义 *由 S 生成的所有向量组成的集合称为 S 的***线性生成集***(linear span), 并记作 $L(S)$.*

换句话说, S 的线性生成集就是令每一个系数 c_i 取遍所有可能的纯量值时所能得到的 S 中向量的全部线性组合所构成的集合. 注意到 $L(S)$ 中向量的线性组合仍在 $L(S)$ 中. 若 $L(S)=\mathbb{R}^n$, 我们称 S 生成空间 $\mathbb{R}^n$.

例 1 令 $S=\{\boldsymbol{A}_1\}$, 则 $L(S)$ 包含 A_1 的所有纯量倍数. □

例 2 当所有纯量 $c_i(1\leqslant i\leqslant k)$ 都为零时, 线性组合 $\sum\limits_{i=1}^{k}c_i\boldsymbol{A}_i$ 等于零向量 $\boldsymbol{O}$, 所以所有的非空集都生成 $\boldsymbol{O}$. 这种所有系数都为零的 $\boldsymbol{O}$ 的表示也称为零向量的**平凡表示**. 另一方面, 也可能有 S 的非平凡的线性组合表示 $\boldsymbol{O}$. 例如, 若 S 中有一个向量是另一个的纯量倍数, 不妨设 $A_2=2A_1$, 则 $\boldsymbol{O}$ 就有很多非平凡表示, 如对任意纯量 t, 有

$$2t\boldsymbol{A}_1-t\boldsymbol{A}_2+0\boldsymbol{A}_3+\cdots+0\boldsymbol{A}_k=\boldsymbol{O}.$$ □

我们对只能用一种方法生成向量的集合 S 特别感兴趣.

定义 *令 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 为 $\mathbb{R}^n$ 上向量的集合, 且它生成向量 $\boldsymbol{X}$, 如果由 $\boldsymbol{X}=\sum\limits_{i=1}^{k}c_i\boldsymbol{A}_i$ 和 $\boldsymbol{X}=\sum\limits_{i=1}^{k}d_i\boldsymbol{A}_i$ 可得对所有 i, 都有*

$$c_i=d_i,\tag{1.10}$$

*则称 S***唯一生成** $\boldsymbol{X}$.

在式 (1.10) 中, 我们假定两个求和运算中的向量 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k$ 的顺序相同且为预先给定的.

定理 1.7 当且仅当集合 S 唯一生成零向量时, 它唯一生成 $L(S)$ 中的所有向量.

证明 若 S 唯一生成 $L(S)$ 中的所有向量, 则它显然也唯一生成 $\boldsymbol{O}$. 现在证明逆命题. 假定 S 唯一生成 $\boldsymbol{O}$, 取 $L(S)$ 上任意向量 $\boldsymbol{X}$, 若 S 有两种方式生成 $\boldsymbol{X}$, 设

$$\boldsymbol{X}=\sum_{i=1}^{k}c_i\boldsymbol{A}_i,\quad \boldsymbol{X}=\sum_{i=1}^{k}d_i\boldsymbol{A}_i,$$

将以上两式相减, 我们有 $\boldsymbol{O}=\sum_{i=1}^{k}(c_i-d_i)\boldsymbol{A}_i$. 另一方面, S 唯一生成 $\boldsymbol{O}$, 故对所有的 $1\leqslant i\leqslant k$, c_i-d_i 都为零, 所以 $c_i=d_i$, 即 S 唯一生成 $\boldsymbol{X}$. □

1.13 线性无关

定理 1.7 体现了唯一生成零向量的集合的重要性, 因此, 我们给这样的集合一个特殊的名字.

定义 如果 $\mathbb{R}^n$ 中的向量组 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 唯一生成零向量, 则称 S 为**线性无关的**, 反之, 则称 S 为**线性相关的**.

也就是说, 线性无关就意味着 S 只能平凡地表示 $\boldsymbol{O}$, 即由 $\sum_{i=1}^{k}c_i\boldsymbol{A}_i=\boldsymbol{O}$ 得所有 $c_i=0$. 而线性相关意味着 S 可以非平凡地表示 $\boldsymbol{O}$, 即存在不全为零的纯量 $c_1,\cdots,c_k$, 使 $\sum_{i=1}^{k}c_i\boldsymbol{A}_i=\boldsymbol{O}$.

尽管线性相关和线性无关是向量组的性质, 我们也经常将这些术语用于向量本身. 例如, 线性无关组中的向量也称为**线性无关向量**. 另外, 我们约定空集为线性无关的.

下面的例子进一步解释线性相关和线性无关.

例 1 若集合 S 的一个子集 T 是线性相关的, 则 S 本身也是线性相关的. 这是因为如果 T 可以非平凡地表示 $\boldsymbol{O}$, 那么 S 也可以非平凡地表示 $\boldsymbol{O}$. 这个命题与命题“线性无关组的所有子集必线性无关”逻辑等价.

例 2 $\mathbb{R}^n$ 中的 n 个单位坐标向量唯一生成 $\boldsymbol{O}$, 所以它们是线性无关的.

例 3 所有包含零向量的向量组是线性相关的. 这是因为若我们用 1 乘零向量, 并用 0 乘所有其余的向量, 那么我们得到 $\boldsymbol{O}$ 的一个非平凡表示.

例 4 在 $\mathbb{R}^2$ 中, 向量组 $S=\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{i}+\boldsymbol{j}\}$ 是线性相关的. 这是因为零向量有如下所示的非平凡表示:

$$\boldsymbol{O}=\boldsymbol{i}+\boldsymbol{j}+(-1)(\boldsymbol{i}+\boldsymbol{j}).$$

在上例中, 我们看到由 i 和 j 生成的空间的某三个向量是线性相关的. 下面的定理表明, 两个向量生成的空间中的任意三个向量都是线性相关的.

定理 1.8　令 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 为 k 个 $\mathbb{R}^n$ 上向量组成的线性无关组, 并令 $L(S)$ 为 S 的线性生成集, 则 $L(S)$ 中的任意 $k+1$ 个向量都是线性相关的.

证明　对 S 中向量个数 k 用归纳法. 首先, 令 $k=1$, 则 S 中只包含一个向量 $\boldsymbol{A}_1$, 且由于 S 线性无关, 所以 $\boldsymbol{A}_1$ 不为零. 任取 $L(S)$ 中两个向量 $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, 则它们都是 $\boldsymbol{A}_1$ 的纯量倍数, 设 $\boldsymbol{B}_1=c_1\boldsymbol{A}_1$, $\boldsymbol{B}_2=c_2\boldsymbol{A}_1$, 则 c_1, c_2 不全为零. 用 c_2 乘 $\boldsymbol{B}_1$, c_1 乘 B_2, 并相减可得

$$c_2\boldsymbol{B}_1-c_1\boldsymbol{B}_2=\boldsymbol{O},$$

这是 $\boldsymbol{O}$ 的一个非平凡表示, 所以 $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ 是线性相关的, 故当 $k=1$ 时, 定理得证.

假设对于 $k-1$ 定理成立, 现证明对于 k 定理也成立. 取 $L(S)$ 中任意 $k+1$ 个向量组成集合 T, 设 $T=\{\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_{k+1}\}$, 我们将证明 T 是线性相关的. 考虑到所有的 $\boldsymbol{B}_i$ 都是 $L(S)$ 中的向量, 我们可将它们写成 S 中元素的线性组合, 设对所有的 $i=1,2,\cdots,k+1$, 有

$$\boldsymbol{B}_i=\sum_{j=1}^{k}a_{ij}\boldsymbol{A}_j. \tag{1.11}$$

观察所有 A_1 的系数 a_{i1}, 并根据它们是否为零分情况讨论.

情形 1　对所有的 $i=1,2,\cdots,k+1$, $a_{i1}=0$. 在这种情况下, 式 (1.11) 中的和与 A_1 无关, 所以所有的 $\boldsymbol{B}_i$ 都是集合 $S'=\{\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k\}$ 的线性生成集中的向量. 又因为 S' 是线性无关的且包含 $k-1$ 个向量, 由归纳假设可知定理在 $k-1$ 时成立, 所以 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_k$ 这 k 个 $L(S')$ 中的向量是线性相关的, 故 T 是线性相关的, 即在这种情况下, 定理成立.

情形 2　a_{i1} 不全为零, 不妨设 $a_{11}\neq 0$(若有必要, 我们总可对 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_{k+1}$ 重新编号使得 $a_{11}\neq 0$), 在式 (1.11) 中令 $i=1$, 并对等号两边同时乘以纯量 $c_i=a_{i1}/a_{11}$, 则有

$$c_i\boldsymbol{B}_1=a_{i1}\boldsymbol{A}_1+\sum_{j=2}^{k}c_ia_{1j}\boldsymbol{A}_j,$$

用上式减式 (1.11), 则可消去 $\boldsymbol{A}_1$ 项, 所以对 $i=2,\cdots,k+1$, 有

$$c_i\boldsymbol{B}_1-\boldsymbol{B}_i=\sum_{j=2}^{k}(c_ia_{1j}-a_{ij})\boldsymbol{A}_j.$$

上式表明对 $2\leqslant i\leqslant k+1$, k 个向量 $c_i\boldsymbol{B}_1-\boldsymbol{B}_i$ 都是 $k-1$ 个线性无关的向量 $\boldsymbol{A}_2$, $\cdots$, $\boldsymbol{A}_k$ 的线性组合. 由归纳假设可知, 这 k 个向量 $c_i\boldsymbol{B}_1-\boldsymbol{B}_i$ 必线性相关, 所以存在不全为零的纯量 $t_2,\cdots,t_{k+1}$, 使得

$$\sum_{i=2}^{k+1} t_i(c_i\boldsymbol{B}_1-\boldsymbol{B}_i)=\boldsymbol{O},$$

由此可得

$$\left(\sum_{i=2}^{k+1} t_i c_i\right)\boldsymbol{B}_1-\sum_{i=2}^{k+1} t_i\boldsymbol{B}_i=\boldsymbol{O}.$$

而这是 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_{k+1}$ 的一个表示零向量的非平凡的线性组合, 所以 $\boldsymbol{B}_1,\cdots,\boldsymbol{B}_{k+1}$ 必线性相关, 故定理得证. □

接下来我们将建立正交性和线性无关性之间的关系.

定义 设 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 为 $\mathbb{R}^n$ 中的向量组, 若对所有的 $\boldsymbol{A}_i\neq\boldsymbol{A}_j$, 都有 $\boldsymbol{A}_i\cdot\boldsymbol{A}_j=0$, 则称 S 为**正交组**, 即正交组中的任意两个不同的向量相互垂直.

定理 1.9 $\mathbb{R}^n$ 中所有非零向量组成的正交组 $S=\{\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k\}$ 都线性无关. 更进一步, 如果 S 生成向量 $\boldsymbol{X}$, 例如

$$\boldsymbol{X}=\sum_{i=1}^{k} c_i\boldsymbol{A}_i, \tag{1.12}$$

则纯量系数 $c_1,\cdots,c_k$ 由如下公式确定:

$$c_j=\frac{\boldsymbol{X}\cdot\boldsymbol{A}_j}{\boldsymbol{A}_j\cdot\boldsymbol{A}_j},\quad j=1,2,\cdots,k. \tag{1.13}$$

注 若 $\boldsymbol{A}_j$ 的范数为 1, 则上式简化为 $c_j=\boldsymbol{X}\cdot\boldsymbol{A}_j$.

证明 首先我们证明 S 是线性无关的. 假定 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_k$ 的一个线性组合为零, 例如 $\sum\limits_{i=1}^{k} c_i\boldsymbol{A}_i=\boldsymbol{O}$, 用 $\boldsymbol{A}_1$ 点乘此式两边, 因为对任意 $i\neq 1$, 都有 $\boldsymbol{A}_1\cdot\boldsymbol{A}_i=0$, 所以有 $c_1(\boldsymbol{A}_1\cdot\boldsymbol{A}_1)=0$. 又因为 $\boldsymbol{A}_1$ 不为零, 所以 $\boldsymbol{A}_1\cdot\boldsymbol{A}_1\neq 0$, 故 $c_1=0$. 同理, 对 $j=2,\cdots,k$, 我们有 $c_j=0$, 所以 S 唯一生成 $\boldsymbol{O}$, 故 S 线性无关.

设 S 如式 (1.12) 所示生成向量 $\boldsymbol{X}$, 对 $j=1,\cdots,k$, 用 $\boldsymbol{A}_j$ 点乘式 (1.12) 的两边, 我们有 $c_j(\boldsymbol{A}_j\cdot\boldsymbol{A}_j)=\boldsymbol{X}\cdot\boldsymbol{A}_j$, 由此可得式 (1.13) 成立. □

定义 若一个正交向量组的所有向量的范数均为 1, 则称它为**标准正交组** (orthonormal set).

单位坐标向量 $\boldsymbol{I}_1,\cdots,\boldsymbol{I}_n$ 就给出一个标准正交组.

1.14 基

由前几节的讨论我们知道人们需要研究唯一生成 $\mathbb{R}^n$ 中所有向量的向量组, 我们把这样的向量组称为**基** (basis).

定义 设 S 为 $\mathbb{R}^n$ 中有限个向量的集合, 若 S 是线性无关的且生成 $\mathbb{R}^n$, 则我们称 S 为 $\mathbb{R}^n$ 的一组基. 更进一步, 如果 S 是正交组, 则称它为 $\mathbb{R}^n$ 的一组**正交基**.

所以 $\mathbb{R}^n$ 的基唯一生成 $\mathbb{R}^n$ 中的所有向量. 单位坐标向量组成的集合就是一组基, 这组基也是一组正交基. 现在我们证明 $\mathbb{R}^n$ 的所有基都恰好包含 n 个向量.

定理 1.10 $\mathbb{R}^n$ 的基具有如下性质:

(a) $\mathbb{R}^n$ 的所有基都恰好有 n 个向量;

(b) $\mathbb{R}^n$ 中的任意一个线性无关的向量组都是 $\mathbb{R}^n$ 的某一组基的子集;

(c) $\mathbb{R}^n$ 中的任意 n 个线性无关的向量组成的集合都是 $\mathbb{R}^n$ 的一组基.

证明 我们知道单位坐标向量 $\boldsymbol{I}_1, \cdots, \boldsymbol{I}_n$ 构成 $\mathbb{R}^n$ 的一组基, 为了证明 (a), 我们将证明 $\mathbb{R}^n$ 的任意两组基中向量的个数相等.

令 S 和 T 为 $\mathbb{R}^n$ 的两组基, 其中 S 有 k 个向量, T 有 r 个向量. 若 $r > k$, 则 T 包含 $L(S)$ 中至少 $k+1$ 个向量, 故由定理 1.8, T 必定线性相关, 这与 T 是 $\mathbb{R}^n$ 的一组基矛盾, 所以 $r \leqslant k$. 由 S 和 T 的对称性可知 $k \leqslant r$, 所以 $k = r$, 即 (a) 得证.

现在我们来证 (b). 令 $S = \{\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k\}$ 为 $\mathbb{R}^n$ 中任意一个线性无关向量组, 若 $L(S) = \mathbb{R}^n$, 则 S 就是一组基. 否则, $\mathbb{R}^n$ 中有某个向量 X 不属于 $L(S)$, 设这个向量与 S 的并集为 $S' = \{\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k, \boldsymbol{X}\}$. 我们来证明 S' 线性无关.

若 S' 线性相关, 则存在不全为零的纯量 $c_1, \cdots, c_{k+1}$, 使得

$$\sum_{i=1}^{k} c_i \boldsymbol{A}_i + c_{k+1}\boldsymbol{X} = \boldsymbol{O}.$$

但是因为 $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k$ 线性无关, 所以 $c_{k+1} \neq 0$, 此时我们可由上式解出 $\boldsymbol{X}$ 并知道它可由 S 线性生成, 这与 $\boldsymbol{X}$ 不属于 $L(S)$ 矛盾, 所以 S' 线性无关并包含 $k+1$ 个向量. 若 S' 是 $\mathbb{R}^n$ 的一组基, 则显然 S 是 S' 的子集, 所以 (b) 得证. 若 S' 不是 $\mathbb{R}^n$ 的一组基, 则重复上述过程, 可由 S' 得到新集合 S'', 它包含 $k+2$ 个向量且线性无关. 若 S'' 是 $\mathbb{R}^n$ 的一组基, 则 (b) 得证. 否则我们继续重复上述过程, 在有限步内, 我们总会得到 $\mathbb{R}^n$ 的一组基, 不然的话我们将得到 $\mathbb{R}^n$ 上 $n+1$ 个线性无关的向量, 这与定理 1.8 矛盾, 因此 (b) 得证.

最后, 我们用 (a) 和 (b) 证明 (c). 令 S 为任意包含 n 个向量的线性无关组, 由 (b), S 是 $\mathbb{R}^n$ 的一组基的子集, 设这组基为 B. 又由 (a), B 恰有 n 个元素, 所以 $S = B$. □

1.15 习 题

1. 令 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 分别表示 $\mathbb{R}^2$ 中的两个单位坐标向量, 求使得向量 $x(\boldsymbol{i}-\boldsymbol{j})+y(\boldsymbol{i}+\boldsymbol{j})$ 分别等于下列各向量的纯量 x 和 y:

   (a) $\boldsymbol{i}$ (b) $\boldsymbol{j}$

   (c) $3\boldsymbol{i}-5\boldsymbol{j}$ (d) $7\boldsymbol{i}+5\boldsymbol{j}$.

2. 设 $\boldsymbol{A} = (1,2)$, $\boldsymbol{B} = (2,-4)$, $\boldsymbol{C} = (2,-3)$ 为 $\mathbb{R}^2$ 中的三个向量, 求纯量 x 和 y 使得 $\boldsymbol{C} = x\boldsymbol{A} + y\boldsymbol{B}$. 有多少对这样的 x, y?

3. 设 $\boldsymbol{A}=(2,-1,1)$, $\boldsymbol{B}=(1,2,-1)$, $\boldsymbol{C}=(2,-11,7)$ 为 $\mathbb{R}^3$ 中的三个向量, 求纯量 x 和 y 使得 $\boldsymbol{C}=x\boldsymbol{A}+y\boldsymbol{B}$.
4. 证明: 若习题 3 中 $\boldsymbol{C}=(2,11,7)$, 则此题无解.
5. 令 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $\mathbb{R}^n$ 中的两个非零向量.
   (a) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 平行, 证明 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关.
   (b) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 不平行, 证明 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性无关.
6. 设 (a,b), (c,d) 为 $\mathbb{R}^2$ 中两个向量, 证明当且仅当 $ad-bc\neq 0$ 时, 它们线性无关.
7. 求所有实数 t, 使得 $\mathbb{R}^2$ 中两个向量 $(1+t,1-t)$ 和 $(1-t,1+t)$ 线性无关.
8. 令 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 为 $\mathbb{R}^3$ 中的单位坐标向量, 证明 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}, \boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}$ 这四个向量线性相关, 但其中任意三个都线性无关.
9. 令 $\boldsymbol{i}, \boldsymbol{j}$ 为 $\mathbb{R}^2$ 中的单位坐标向量, 并令 $S=\{\boldsymbol{i}, \boldsymbol{i}+\boldsymbol{j}\}$,
   (a) 证明 S 是线性无关组.
   (b) 证明 $\boldsymbol{j}$ 属于 S 的线性生成集.
   (c) 将 $3\boldsymbol{i}-4\boldsymbol{j}$ 表示为 $\boldsymbol{i}$ 和 $\boldsymbol{i}+\boldsymbol{j}$ 的线性组合.
   (d) 证明 $L(S)=\mathbb{R}^2$.
10. 考虑 $\mathbb{R}^3$ 中三个向量 $\boldsymbol{A}=\boldsymbol{i}$, $\boldsymbol{B}=\boldsymbol{i}+\boldsymbol{j}$, $\boldsymbol{C}=\boldsymbol{i}+\boldsymbol{j}+3\boldsymbol{k}$.
    (a) 证明向量组 $\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\}$ 线性无关.
    (b) 将 $\boldsymbol{j}$ 和 $\boldsymbol{k}$ 表示为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的线性组合.
    (c) 将 $2\boldsymbol{i}-3\boldsymbol{j}+5\boldsymbol{k}$ 表示为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的线性组合.
    (d) 证明 $\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\}$ 为 $\mathbb{R}^3$ 的一组基.
11. 令 $\boldsymbol{A}=(1,2)$, $\boldsymbol{B}=(2,-4)$, $\boldsymbol{C}=(2,-3)$, $\boldsymbol{D}=(1,-2)$ 为 $\mathbb{R}^2$ 中的四个向量, 列出 $\{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}\}$ 的所有线性无关的非空子集.
12. 令 $\boldsymbol{A}=(1,1,1,0)$, $\boldsymbol{B}=(0,1,1,1)$, $\boldsymbol{C}=(1,1,0,0)$ 为 $\mathbb{R}^4$ 中的三个向量,
    (a) $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 是线性无关的还是线性相关的?
    (b) 举出一个非零向量 $\boldsymbol{D}$ 的例子, 使得 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 线性相关.
    (c) 举出一个向量 $\boldsymbol{E}$ 的例子, 使得 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{E}$ 线性无关.
    (d) 将 $\boldsymbol{X}=(1,2,3,4)$ 表示为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{E}$ 的线性组合, 其中 $\boldsymbol{E}$ 为 (c) 中所举的例子.
13. (a) 证明 $\mathbb{R}^3$ 中向量 $(\sqrt{3},1,0)$, $(1,\sqrt{3},1)$, $(0,1,\sqrt{3})$ 线性无关.
    (b) 证明如下三个向量线性相关: $(\sqrt{2},1,0)$, $(1,\sqrt{2},1)$, $(0,1,\sqrt{2})$.
    (c) 求使得如下三个向量线性相关的所有实数 t: $(t,1,0)$, $(1,t,1)$, $(0,1,t)$.
14. 对下面 $\mathbb{R}^4$ 中各向量组分别找出一个包含向量个数最多的线性无关的部分组:
    (a) $\{(1,0,1,0),(1,1,1,1),(0,1,0,1),(2,0,-1,0)\}$
    (b) $\{(1,1,1,1),(1,-1,1,1),(1,-1,-1,1),(1,-1,-1,-1)\}$
    (c) $\{(1,1,1,1),(0,1,1,1),(0,0,1,1),(0,0,0,1)\}$
15. 给定 $\mathbb{R}^n$ 中三个线性无关的向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, 证明或推翻下述各命题:
    (a) $\boldsymbol{A}+\boldsymbol{B}$, $\boldsymbol{B}+\boldsymbol{C}$, $\boldsymbol{A}+\boldsymbol{C}$ 线性无关;
    (b) $\boldsymbol{A}-\boldsymbol{B}$, $\boldsymbol{B}+\boldsymbol{C}$, $\boldsymbol{A}+\boldsymbol{C}$ 线性无关.
16. (a) 证明: 若向量组 S 包含 $\mathbb{R}^3$ 中三个向量, 则当且仅当它的线性生成集 $L(S)$ 包含三个单位坐标向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 时, 它是 $\mathbb{R}^3$ 的一组基.
    (b) 将 (a) 的结论推广到 $\mathbb{R}^n$ 上并加以证明.
17. 找出 $\mathbb{R}^3$ 中包含 $(0,1,1)$ 和 $(1,1,1)$ 的不同的两组基.
18. 找出 $\mathbb{R}^4$ 中交集为 $\{(0,1,1,1),(1,1,1,1)\}$ 的两组基.
19. 考虑 $\mathbb{R}^3$ 中如下三个向量组:

$S = \{(1,1,1),(0,1,2),(1,0,-1)\}$, $T = \{(2,1,0),(2,0,-2)\}$, $U = \{\{(1,2,3),(1,3,5)\}$,

(a) 证明 $L(T) \subseteq L(S)$.

(b) 找出集合 $L(S)$, $L(T)$, $L(U)$ 之间的所有包含关系.

20. 令 A 和 B 表示 $\mathbb{R}^n$ 的两个有限子集, 并令 $L(A)$ 和 $L(B)$ 表示它们的线性生成集, 证明下述各命题.

(a) 若 $A \subseteq B$, 则 $L(A) \subseteq L(B)$.

(b) $L(A \cap B) \subseteq L(A) \cap L(B)$.

(c) 给出 $L(A \cap B) \neq L(A) \cap L(B)$ 的一个例子.

1.16 复数的 n 元组构成的向量空间 $\mathbb{C}^n$

在 1.2 节中, 我们定义向量空间 $\mathbb{R}^n$ 为全体实 n 元组的集合. 而向量的相等、加法和数乘运算如下所示, 都是用分量的运算定义的: 设 $\boldsymbol{A} = (a_1, \cdots, a_n)$, $\boldsymbol{B} = (b_1, \cdots, b_n)$, 则

$$\boldsymbol{A} = \boldsymbol{B} \text{ 等价于对所有的 } i = 1, 2, \cdots, n, a_i = b_i,$$

$$\boldsymbol{A} + \boldsymbol{B} = (a_1 + b_1, \cdots, a_n + b_n), \quad c\boldsymbol{A} = (ca_1, \cdots, ca_n).$$

如果上式中的所有纯量 a_i, b_i, c 都为复数, 那么我们得到的新代数系统称为**复 n 维空间** (complex n-space), 记作 $\mathbb{C}^n$, 其中 $\mathbb{C}$ 表示该空间中的纯量都是复数.

因为实数和复数具有相同的加法、乘法、除法性质, 本章中讨论过的所有只用到向量的加法和数乘的定理在 $\mathbb{C}^n$ 中仍然成立, 区别仅在于这里参与运算的所有纯量均为复数.

我们做这样的推广并不仅仅是为了把 n 维空间的概念一般化. 相反, 复向量空间的概念是因为在线性微分方程理论和现代量子力学中有需要而自然引入的, 所以复向量空间的研究就变得相当重要. 尽管许多关于 $\mathbb{R}^n$ 的定理都可原封不动地移植到 $\mathbb{C}^n$ 上来, 但是那些与点积有关的定理仍然需要作适当的改动才能用在 $\mathbb{C}^n$ 上. 我们将用不全为零的实数的平方和为正这一事实来证明非零向量与自己的点积为正. 由于复数的平方和可能为负, 为了保留正定性, 我们必须修改点积的定义. 在 $\mathbb{C}^n$ 中, 我们用如下关系定义点积.

定义 设 $\boldsymbol{A} = (a_1, \cdots, a_n)$, $\boldsymbol{B} = (b_1, \cdots, b_n)$ 为 $\mathbb{C}^n$ 中的两个向量, 我们用如下公式定义它们的点积 $\boldsymbol{A} \cdot \boldsymbol{B}$:

$$\boldsymbol{A} \cdot \boldsymbol{B} = \sum_{k=1}^{n} a_k \overline{b_k},$$

其中 $\overline{b_k}$ 为 b_k 的共轭复数.

这个定义与我们前面给出的 $\mathbb{R}^n$ 中的点积的定义一致, 这是因为当 b_k 为实数时, $\overline{b_k} = b_k$. 定理 1.2 中所述的点积的基本性质将变成如下形式.

定理 1.11 对任意 $\mathbb{C}^n$ 上的向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 以及任意复纯量 c, 我们有

(a) $\boldsymbol{A} \cdot \boldsymbol{B} = \overline{\boldsymbol{B} \cdot \boldsymbol{A}}$,

(b) $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\cdot\boldsymbol{B}+\boldsymbol{A}\cdot\boldsymbol{C}$,

(c) $c(\boldsymbol{A}\cdot\boldsymbol{B})=(c\boldsymbol{A})\cdot\boldsymbol{B}=\boldsymbol{A}\cdot(\bar{c}\boldsymbol{B})$,

(d) $\boldsymbol{A}\cdot\boldsymbol{A}>0$, 若 $\boldsymbol{A}\neq\boldsymbol{O}$,

(e) $\boldsymbol{A}\cdot\boldsymbol{A}=0$, 若 $\boldsymbol{A}=\boldsymbol{O}$. □

这些性质都可由定义直接得到, 我们将它们的证明留给读者作为练习. 读者必须注意到在性质 (a) 中, 当交换点积运算中两个向量的次序时, 我们对结果取共轭. 同样在性质 (c) 中, 当纯量 c 从点乘号的一边移到另一边时, 也需对 c 取共轭.

Cauchy-Schwarz 不等式在 $\mathbb{C}^n$ 中具有如下形式:

$$|\boldsymbol{A}\cdot\boldsymbol{B}|^2\leqslant(\boldsymbol{A}\cdot\boldsymbol{A})(\boldsymbol{B}\cdot\boldsymbol{B}). \tag{1.14}$$

上式的证明与定理 1.3 的证明类似. 考虑向量 $\boldsymbol{C}=x\boldsymbol{A}-y\boldsymbol{B}$, 其中 $x=\boldsymbol{B}\cdot\boldsymbol{B}$, $y=\boldsymbol{A}\cdot\boldsymbol{B}$, 并计算 $\boldsymbol{C}\cdot\boldsymbol{C}$, 由不等式 $\boldsymbol{C}\cdot\boldsymbol{C}\geqslant 0$ 即可得式 (1.14). 证明的细节留给读者作为练习.

因为一个向量和它自己的点积非负, 我们可用和 $\mathbb{R}^n$ 中相同的公式定义 $\mathbb{C}^n$ 上向量的范数:

$$\|\boldsymbol{A}\|=(\boldsymbol{A}\cdot\boldsymbol{A})^{1/2}.$$

定理 1.4 中范数的基本性质也可原封不动地移植到 $\mathbb{C}^n$ 中来. 我们特别指出三角不等式 $\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$ 在 $\mathbb{C}^n$ 中也成立.

$\mathbb{C}^n$ 中向量的正交由关系 $\boldsymbol{A}\cdot\boldsymbol{B}=0$ 定义. 和实空间一样, 当且仅当 $\mathbb{C}^n$ 中两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 满足 Pythagoras 恒等式

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2$$

时它们正交.

在 $\mathbb{C}^n$ 中的线性生成集、线性无关、线性相关、基的定义都和 $\mathbb{R}^n$ 中的完全相同, 定理 1.7~1.10 及其证明在 $\mathbb{C}^n$ 中也成立.

1.17 习 题

1. 令 $\boldsymbol{A}=(1,\mathrm{i})$, $\boldsymbol{B}=(\mathrm{i},-\mathrm{i})$, $\boldsymbol{C}=(2\mathrm{i},1)$ 为 $\mathbb{C}^2$ 中三个向量, 计算以下各点积:

   (a) $\boldsymbol{A}\cdot\boldsymbol{B}$　　(b) $\boldsymbol{B}\cdot\boldsymbol{A}$

   (c) $(\mathrm{i}\boldsymbol{A})\cdot\boldsymbol{B}$　　(d) $\boldsymbol{A}\cdot(\mathrm{i}\boldsymbol{B})$

   (e) $(\mathrm{i}\boldsymbol{A})\cdot(\mathrm{i}\boldsymbol{B})$　　(f) $\boldsymbol{B}\cdot\boldsymbol{C}$

   (g) $\boldsymbol{A}\cdot\boldsymbol{C}$　　(h) $(\boldsymbol{B}+\boldsymbol{C})\cdot\boldsymbol{A}$

   (i) $(\boldsymbol{A}-\boldsymbol{C})\cdot\boldsymbol{B}$　　(j) $(\boldsymbol{A}-\mathrm{i}\boldsymbol{B})\cdot(\boldsymbol{A}+\mathrm{i}\boldsymbol{B})$

2. 设 $\boldsymbol{A}=(2,1,-\mathrm{i})$, $\boldsymbol{B}=(\mathrm{i},-1,2\mathrm{i})$, 求 $\mathbb{C}^3$ 中一个与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都正交的非零向量 $\boldsymbol{C}$.

3. 证明: 对 $\mathbb{C}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 有恒等式

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2+\boldsymbol{A}\cdot\boldsymbol{B}+\overline{\boldsymbol{A}\cdot\boldsymbol{B}}.$$

4. 证明: 对 $\mathbb{C}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 有恒等式

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2-\|\boldsymbol{A}-\boldsymbol{B}\|^2=2(\boldsymbol{A}\cdot\boldsymbol{B}+\overline{\boldsymbol{A}\cdot\boldsymbol{B}}).$$

5. 证明：对 $\mathbb{C}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 有恒等式

$$\|\boldsymbol{A}+\boldsymbol{B}\|^2+\|\boldsymbol{A}-\boldsymbol{B}\|^2=2\|\boldsymbol{A}\|^2+2\|\boldsymbol{B}\|^2.$$

6. (a) 证明：对 $\mathbb{C}^n$ 中任意两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, $\boldsymbol{A}\cdot\boldsymbol{B}+\overline{\boldsymbol{A}\cdot\boldsymbol{B}}$ 是实数.

   (b) 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $\mathbb{C}^n$ 中两个非零向量, 证明

$$-2\leqslant\frac{\boldsymbol{A}\cdot\boldsymbol{B}+\overline{\boldsymbol{A}\cdot\boldsymbol{B}}}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|}\leqslant 2$$

7. 定义 $\mathbb{C}^n$ 中两个非零向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的夹角 θ 为

$$\theta=\arccos\frac{\frac{1}{2}(\boldsymbol{A}\cdot\boldsymbol{B}+\overline{\boldsymbol{A}\cdot\boldsymbol{B}})}{\|\boldsymbol{A}\|\|\boldsymbol{B}\|},$$

   习题 6 中的不等式表明在区间 $[0,\pi]$ 上存在唯一的一个角 θ 满足上式, 证明

$$\|\boldsymbol{A}-\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2-2\|\boldsymbol{A}\|\|\boldsymbol{B}\|\cos\theta.$$

8. 用习题 7 给出的定义计算 $\mathbb{C}^5$ 中两个向量 $\boldsymbol{A}=(1,0,\mathrm{i},\mathrm{i},\mathrm{i})$ 和 $\boldsymbol{B}=(\mathrm{i},\mathrm{i},\mathrm{i},0,\mathrm{i})$ 的夹角.

9. (a) 证明向量 $\boldsymbol{A}=(1,0,0)$, $\boldsymbol{B}=(0,\mathrm{i},0)$, $\boldsymbol{C}=(1,1,\mathrm{i})$ 构成 $\mathbb{C}^3$ 的一组基.

   (b) 将向量 $(5,2-\mathrm{i},2\mathrm{i})$ 表示为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的线性组合.

10. 单位坐标向量 $\boldsymbol{I}_1,\cdots,\boldsymbol{I}_n$ 构成 $\mathbb{R}^n$ 的一组基, 证明它们同时也构成 $\mathbb{C}^n$ 的一组基.

第 2 章　向量代数在解析几何中的应用

2.1　引　　言

本章将讨论向量代数在直线、平面和二次曲线的研究中的应用. 我们熟悉的二维和三维空间中直线和平面的性质都将推广到 n 维空间中.

几何学最早由 Euclid 在公元前约 300 年作为一种演绎系统引入. 他先引入一组描述点和直线性质的公理 (公设). Euclid 将**点** (point) 和**直线** (line) 看作不加定义的原始概念, 其他的概念都是用点和直线定义的, 而定理则由公理系统地推导而得. Euclid 列出了十条公理并试图由此推导出他想获得的所有定理, 然而人们已经证明了这些公理并不足以推导出整个欧氏几何理论. 例如, 在他的第一个定理的证明中, Euclid 作了一个隐含的关于两圆相交的假设, 而这个假设并不在任何一条公理中.

在这之后, 人们又给出了其他确实能证明欧氏几何所有定理的公理系统, 其中最著名的是由德国数学家 David Hilbert(1862–1943) 提出的公理系统. 在他 1899 年出版的如今已成为经典的著作《几何基础》(*Grundlagen der Geometrie*, 英译本 *The Foundations of Geometry*, Open Court Publishing Co., 1947) 中, Hilbert 提出了他的公理系统. 这本书奠定了 20 世纪抽象数学的基础, 这本书在 Hilbert 生前一共出了七个德语版本.

在平面几何中, Hilbert 先给出了五个未定义的概念：点、直线、**在 ⋯ 之上** (on) (点和直线间的一种关系)、**在 ⋯ 之间**(between)(一个点和一对点之间的一种关系), **合同**(congruent)(点对之间的一种关系), 然后他列出了十五条公理, 由这些公理可推导出整个欧氏平面几何理论. 同样他给出了二十一条关于六个未定义概念的公理, 并由此建立整个欧氏立体几何理论.

解析几何则略有不同, 我们用实数定义点、线、在 ⋯ 之上、在 ⋯ 之间等概念, 然而我们却并不给出实数的定义. 这样得到的数学结构称为欧氏几何的**解析模型**(analytic model). 在这个模型中, 将 Hilbert 公理作为定理并利用实数的性质证明它们. 本书中, 我们并不给出所有的 Hilbert 公理的描述, 相反, 我们仅仅阐述如何用数来定义几何中的原始概念, 同时我们也给出几个定理的证明来说明解析几何方法.

2.2　n 维空间中的直线

本节中我们用实数定义点、直线、在 $\cdots$ 之上的概念. 这些定义符合我们对二维和三维空间中的对应概念的直观感觉, 同时它们也在所有的 n 维空间中有意义, 其中 $n \geqslant 1$.

我们将 $\mathbb{R}^n$ 中的向量 (实数的 n 元有序组) 定义为**点**. 我们把点和向量看作同一个概念, 不加以区分. 向量空间 $\mathbb{R}^n$ 称为 n 维欧氏空间的解析模型, 或者简单地称为 n **维欧氏空间**. 我们用 $\mathbb{R}^n$ 上的向量加法和纯量乘法来定义直线.

定义　给定点 $\boldsymbol{P}$ 和非零向量 $\boldsymbol{A}$, 由所有的点 $\boldsymbol{P}+t\boldsymbol{A}$ 组成的集合称为过 $\boldsymbol{P}$ 且平行于 $\boldsymbol{A}$ 的直线, 其中 t 取遍所有实数. 我们将这条直线记作 $L(\boldsymbol{P};\boldsymbol{A})$, 即

$$L(\boldsymbol{P};\boldsymbol{A})=\{\boldsymbol{P}+t\boldsymbol{A}:t\in\mathbb{R}\}\quad \text{或者简单地写作}\quad L(\boldsymbol{P};\boldsymbol{A})=\{\boldsymbol{P}+t\boldsymbol{A}\}.$$

若对点 $\boldsymbol{Q}$, 存在纯量 t 满足 $\boldsymbol{Q}=\boldsymbol{P}+t\boldsymbol{A}$, 则称 $\boldsymbol{Q}$ 在直线 $L(\boldsymbol{P};\boldsymbol{A})$ 上.

在记号 $L(\boldsymbol{P};\boldsymbol{A})$ 中, 点 $\boldsymbol{P}$ 写在前面, 由于 $\boldsymbol{P}$ 与 $t=0$ 相对应, 因此点 $\boldsymbol{P}$ 在该直线上. 点 $\boldsymbol{A}$ 称作该直线的**方向向量** (direction vector). 过原点的直线 $L(\boldsymbol{O};\boldsymbol{A})$ 是 $\boldsymbol{A}$ 的线性生成集, 它包含所有 $\boldsymbol{A}$ 的纯量倍数. 过点 $\boldsymbol{P}$ 且平行于 $\boldsymbol{A}$ 的直线可通过将点 $\boldsymbol{P}$ 与 $\boldsymbol{A}$ 的线性生成集中的所有向量相加而得.

图 2.1 表示 $\mathbb{R}^3$ 中直线定义的几何解释. 若我们将点 $\boldsymbol{P}+t\boldsymbol{A}$ 看作以原点为起点的向量 $\boldsymbol{P}+t\boldsymbol{A}$ 的终点, 那么当 t 取遍所有实数时, 点 $\boldsymbol{P}+t\boldsymbol{A}$ 构成的轨迹即为过点 $\boldsymbol{P}$ 且平行于向量 $\boldsymbol{A}$ 的直线. 图 2.1 中也画出了当 t 取某些值时直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{O};\boldsymbol{A})$ 上对应的点.

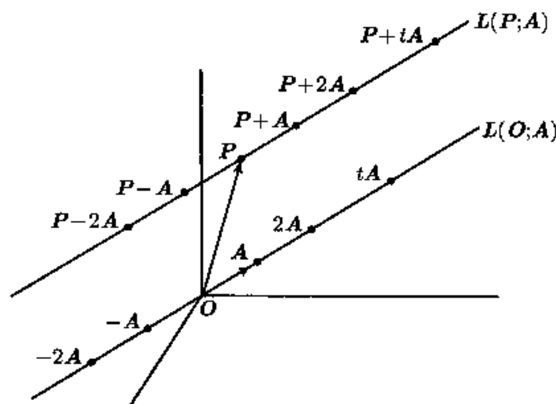


图 2.1　过点 P 且平行于 A 的直线 $L(\boldsymbol{P};\boldsymbol{A})$ 及其与过点 $\boldsymbol{O}$ 且平行于 A 的直线 $L(\boldsymbol{O};\boldsymbol{A})$ 的几何关系

2.3 $\mathbb{R}^n$ 中直线的一些简单性质

首先, 我们证明在 $L(\boldsymbol{P};\boldsymbol{A})$ 的定义中方向向量 $\boldsymbol{A}$ 可用平行于 $\boldsymbol{A}$ 的任意一个向量代替, 而所得的直线不变 (记住当存在非零纯量 c 使得 $\boldsymbol{A}=c\boldsymbol{B}$ 时, 称向量 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 平行).

定理 2.1 过同一点 $\boldsymbol{P}$ 的两条直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{P};\boldsymbol{B})$ 相同, 当且仅当它们的方向向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 互相平行.

证明 首先设 $L(\boldsymbol{P};\boldsymbol{A})=L(\boldsymbol{P};\boldsymbol{B})$. 取 $L(\boldsymbol{P};\boldsymbol{A})$ 上不同于 $\boldsymbol{P}$ 的点, 不妨设该点为 $\boldsymbol{P}+\boldsymbol{A}$, 那么该点也在 $L(\boldsymbol{P};\boldsymbol{B})$ 上, 所以存在纯量 c 使得 $\boldsymbol{P}+\boldsymbol{A}=\boldsymbol{P}+c\boldsymbol{B}$, 故 $\boldsymbol{A}=c\boldsymbol{B}$. 另外, 由于 $\boldsymbol{A}\neq\boldsymbol{O}$, 所以 $c\neq 0$, 因此 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 互相平行.

再设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 互相平行. 那么存在纯量 $c\neq 0$ 使 $\boldsymbol{A}=c\boldsymbol{B}$. 若点 $\boldsymbol{Q}$ 在 $L(\boldsymbol{P};\boldsymbol{A})$ 上, 则存在纯量 t 使得 $\boldsymbol{Q}=\boldsymbol{P}+t\boldsymbol{A}=\boldsymbol{P}+t(c\boldsymbol{B})=\boldsymbol{P}+(tc)\boldsymbol{B}$, 所以 $\boldsymbol{Q}$ 在 $L(\boldsymbol{P};\boldsymbol{B})$ 上, 因此有 $L(\boldsymbol{P};\boldsymbol{A})\subseteq L(\boldsymbol{P};\boldsymbol{B})$. 同理可得 $L(\boldsymbol{P};\boldsymbol{B})\subseteq L(\boldsymbol{P};\boldsymbol{A})$, 故 $L(\boldsymbol{P};\boldsymbol{A})=L(\boldsymbol{P};\boldsymbol{B})$. □

接下来我们证明在 $L(\boldsymbol{P};\boldsymbol{A})$ 的定义中, 可用该直线上任意点 $\boldsymbol{Q}$ 代替 $\boldsymbol{P}$ 而所得的直线不变.

定理 2.2 对有同一个方向向量 $\boldsymbol{A}$ 的直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{Q};\boldsymbol{A})$, 当且仅当 $\boldsymbol{Q}$ 在直线 $L(\boldsymbol{P};\boldsymbol{A})$ 上时这两条直线相同.

证明 设 $L(\boldsymbol{P};\boldsymbol{A})=L(\boldsymbol{Q};\boldsymbol{A})$, 由于 $\boldsymbol{Q}$ 在 $L(\boldsymbol{Q};\boldsymbol{A})$ 上, 所以它也在 $L(\boldsymbol{P};\boldsymbol{A})$ 上. 反之, 设 $\boldsymbol{Q}$ 在 $L(\boldsymbol{P};\boldsymbol{A})$ 上, 不妨设 $\boldsymbol{Q}=\boldsymbol{P}+c\boldsymbol{A}$, 我们来证明 $L(\boldsymbol{P};\boldsymbol{A})=L(\boldsymbol{Q};\boldsymbol{A})$. 取 $L(\boldsymbol{P};\boldsymbol{A})$ 上任意点 $\boldsymbol{X}$, 则存在纯量 t 使 $\boldsymbol{X}=\boldsymbol{P}+t\boldsymbol{A}$. 又因为 $\boldsymbol{P}=\boldsymbol{Q}-c\boldsymbol{A}$, 所以有 $\boldsymbol{X}=\boldsymbol{Q}-c\boldsymbol{A}+t\boldsymbol{A}=\boldsymbol{Q}+(t-c)\boldsymbol{A}$, 因此 $\boldsymbol{X}$ 也在 $L(\boldsymbol{Q};\boldsymbol{A})$ 上, 故 $L(\boldsymbol{P};\boldsymbol{A})\subseteq L(\boldsymbol{Q};\boldsymbol{A})$. 同理可证 $L(\boldsymbol{Q};\boldsymbol{A})\subseteq L(\boldsymbol{P};\boldsymbol{A})$, 因此这两条直线相同. □

平行公设 (parallel postulate) 是 Euclid 最著名的公设之一. 这个公设和下述命题在逻辑上等价:“过一个给定点, 存在且仅存在一条与给定直线平行的直线.” 我们将证明这个性质是定理 2.1 的一个简单推论. 为此我们需要给出两条直线平行的定义.

定义 对 n 维空间中的两条直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{Q};\boldsymbol{B})$, 如果它们的方向向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相互平行, 则称这两条直线**平行**.

定理 2.3 给定直线 L 和不在 L 上的一点 $\boldsymbol{Q}$, 则存在且仅存在一条包含 $\boldsymbol{Q}$ 且平行于 L 的直线 L'.

证明 设直线 L 的方向向量为 $\boldsymbol{A}$, 考虑直线 $L'=L(\boldsymbol{Q};\boldsymbol{A})$. 显然它包含点 $\boldsymbol{Q}$ 且平行于 L. 定理 2.1 告诉我们这是唯一一条同时具有这两条性质的直线. □

注 在很长的一段时间里, 数学家们认为平行公设可由 Euclid 的其他公设推导而得, 但是所有证明这

个结论的尝试都失败了. 直到 19 世纪早期, 数学家 Carl F. Gauss(1777–1855)、J. Bolyai(1802–1860)、N. I. Lobatchevski(1791–1856) 开始相信无法由其他公设推导出平行公设并且开始建立非欧几何, 即平行公设不成立的几何. 这几位数学家的工作鼓励其他数学家和科学家重新审视那些“公认的真理”, 并向几个世纪以来一直被奉为圣典的公理发起挑战.

同样易证直线的下述性质, Euclid 将该性质作为公理列出.

定理 2.4　*两个相异点决定一条直线. 即若给定点 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, $\boldsymbol{P} \neq \boldsymbol{Q}$, 则有且仅有一条同时包含 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 的直线. 可用集合 $\{\boldsymbol{P} + t(\boldsymbol{Q} - \boldsymbol{P})\}$ 表示这条直线.*

证明　令 L 为过点 $\boldsymbol{P}$ 且平行于 $\boldsymbol{Q} - \boldsymbol{P}$ 的直线, 即令

$$L = L(\boldsymbol{P}; \boldsymbol{Q} - \boldsymbol{P}) = \{\boldsymbol{P} + t(\boldsymbol{Q} - \boldsymbol{P})\}.$$

这条直线同时包含 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$(令 $t = 0$ 可得 L 包含 $\boldsymbol{P}$, 令 $t = 1$ 可得 L 包含 $\boldsymbol{Q}$). 令 L' 为同时包含 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 的任意直线, 我们将证明 $L' = L$. 由于 L' 包含 $\boldsymbol{P}$, 所以存在非零向量 $\boldsymbol{A}$ 使得 $L' = L(\boldsymbol{P}; \boldsymbol{A})$. 又由于 L' 也包含 $\boldsymbol{Q}$, 所以存在纯量 c 使得 $\boldsymbol{P} + c\boldsymbol{A} = \boldsymbol{Q}$, 故有 $\boldsymbol{Q} - \boldsymbol{P} = c\boldsymbol{A}$. 因为 $\boldsymbol{Q} \neq \boldsymbol{P}$, 所以 $c \neq 0$, 因此 $\boldsymbol{Q} - \boldsymbol{P}$ 平行于 $\boldsymbol{A}$, 由定理 2.2 我们有 $L' = L(\boldsymbol{P}; \boldsymbol{A}) = L(\boldsymbol{P}; \boldsymbol{Q} - \boldsymbol{P}) = L$. □

例　定理 2.4 给出了一个判断给定点 $\boldsymbol{Q}$ 是否在给定直线 $L(\boldsymbol{P}; \boldsymbol{A})$ 上的简单方法. 它告诉我们当且仅当 $\boldsymbol{Q} - \boldsymbol{P}$ 平行于 $\boldsymbol{A}$ 时 $\boldsymbol{Q}$ 在 $L(\boldsymbol{P}; \boldsymbol{A})$ 上. 例如, 对三维空间中的直线 $L(\boldsymbol{P}; \boldsymbol{A})$, 其中 $\boldsymbol{P} = (1, 2, 3)$, $\boldsymbol{A} = (2, -1, 5)$. 为了判断 $\boldsymbol{Q} = (1, 1, 4)$ 是否在该直线上, 我们考察 $\boldsymbol{Q} - \boldsymbol{P} = (0, -1, 1)$, 由于 $\boldsymbol{Q} - \boldsymbol{P}$ 不是 $\boldsymbol{A}$ 的纯量倍数, 所以点 $(1, 1, 4)$ 不在该直线上. 另一方面, 如果 $\boldsymbol{Q} = (5, 0, 13)$, 则因为 $\boldsymbol{Q} - \boldsymbol{P} = (4, -2, 10) = 2\boldsymbol{A}$, 所以此时 $\boldsymbol{Q}$ 在该直线上. □

$\mathbb{R}^n$ 中两个向量的线性相关性也可用几何语言描述.

定理 2.5　*对 $\mathbb{R}^n$ 中两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 当且仅当它们在过原点的同一条直线上时, 这两个向量线性相关.*

证明　若 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 为零, 则结论显然成立. 若它们都不为零, 则当且仅当存在纯量 t 使 $\boldsymbol{B} = t\boldsymbol{A}$ 时, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关. 又因为当且仅当 $\boldsymbol{B}$ 在过原点且平行于 $\boldsymbol{A}$ 的直线上时, 存在纯量 t 使 $\boldsymbol{B} = t\boldsymbol{A}$, 所以定理得证. □

2.4　n 维空间中的直线和向量值函数

将直线看作质点的运动轨迹是有益的. 可用一个由下式定义的向量值函数 $\boldsymbol{X}$ 表示质点在时刻 t 的位置:

$$\boldsymbol{X}(t) = \boldsymbol{P} + t\boldsymbol{A}. \tag{2.1}$$

显然, 对所有实数 t, $\boldsymbol{X}$ 都有意义. 这个 $\boldsymbol{X}$ 是单个实变量的向量值函数的一个例子, 它的定义域为所有实数组成的集合, 值域为直线 $L(\boldsymbol{P}; \boldsymbol{A})$.

我们称式 (2.1) 中的纯量 t 为**参数** (parameter), 称式 (2.1) 本身为直线 $L(\boldsymbol{P}; \boldsymbol{A})$ 的**向量参数方程**, 或简单地称为**向量方程**. 如果我们令 t 代表时间, 那么可将 $\boldsymbol{X}(t)$

看作表示在时刻 t 该质点位置的位置向量. 函数观点非常重要, 它不仅为表示直线而且也为表示一般空间曲线提供了一种十分自然的方法.

如果一条直线经过相异的两点 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 那么可用 $\boldsymbol{Q}-\boldsymbol{P}$ 作为它在方程 (2.1) 中的方向向量 $\boldsymbol{A}$, 此时该直线的向量方程化为

$$\boldsymbol{X}(t)=\boldsymbol{P}+t(\boldsymbol{Q}-\boldsymbol{P}) \quad \text{或} \quad \boldsymbol{X}(t)=t\boldsymbol{Q}+(1-t)\boldsymbol{P}.$$

注意对直线 $L(\boldsymbol{P};\boldsymbol{A})$ 上的两点 $\boldsymbol{X}(a)$ 和 $\boldsymbol{X}(b)$, 当且仅当 $\boldsymbol{P}+a\boldsymbol{A}=\boldsymbol{P}+b\boldsymbol{A}$ 时, 这两点相同, 此时 $(a-b)\boldsymbol{A}=\boldsymbol{O}$. 由于 $\boldsymbol{A}$ 非零, 所以当且仅当 $a=b$ 时 $(a-b)\boldsymbol{A}=\boldsymbol{O}$, 因此直线上参数 t 不相等的点互不相同.

现在我们考虑给定直线上互不相同的三个点 $\boldsymbol{X}(a)$, $\boldsymbol{X}(b)$, $\boldsymbol{X}(c)$, 其中 $a<b$. 当 c 在 a 和 b 之间即 $a<c<b$ 时, 我们称 $\boldsymbol{X}(c)$ **在** $\boldsymbol{X}(a)$ **和** $\boldsymbol{X}(b)$**之间** (between).

我们可用范数来定义合同. 设有两对点 $\boldsymbol{P},\boldsymbol{Q}$ 和 $\boldsymbol{P}',\boldsymbol{Q}'$, 若 $\|\boldsymbol{P}-\boldsymbol{Q}\|=\|\boldsymbol{P}'-\boldsymbol{Q}'\|$, 则称点对 $\boldsymbol{P},\boldsymbol{Q}$ 与点对 $\boldsymbol{P}',\boldsymbol{Q}'$ **合同** (congruence). 范数 $\|\boldsymbol{P}-\boldsymbol{Q}\|$ 称为 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 之间的**距离**.

这样一来, 我们给出了 n 维欧氏空间的解析模型中的全部基本概念: 点, 直线, 在 $\cdots$ 之上, 在 $\cdots$ 之间, 合同. 下一节我们将讨论三维空间和二维空间中直线的参数方程.

2.5 三维空间和二维空间中的直线

设式 (2.1) 表示的直线在三维空间中, 接下来我们用分量的形式表示向量方程 (2.1). 令 $\boldsymbol{P}=(p,q,r)$, $\boldsymbol{A}=(a,b,c)$, $\boldsymbol{X}(t)=(x,y,z)$, 则向量方程 (2.1) 等价于下列三个纯量方程:

$$x=p+ta, \quad y=q+tb, \quad z=r+tc. \tag{2.2}$$

我们称这三个方程为该直线的**纯量参数方程**或简单地称它们为该直线的**参数方程**. 在进行涉及分量的运算时, 这些方程非常有用. 不过向量方程更为简单, 而且在研究直线的一般性质时, 它也更为自然.

若我们讨论的是一个二维空间中的直线, 那么我们只需要式 (2.2) 中的前两个方程. 这样一来, 我们就可消去这两个参数方程中的参数 t. 用 b 乘第一个方程, 用 a 乘第二个方程, 然后将它们相减, 我们得到如下关系:

$$b(x-p)-a(y-q)=0, \tag{2.3}$$

这个方程称为该直线的**笛卡儿方程**. 当 $a\neq 0$ 时, 上式可写成**点斜式** (point-slope form) 方程:

$$y-q=\frac{b}{a}(x-p),$$

其中, 点 (p,q) 在该直线上, b/a 为该直线的斜率.

笛卡儿方程 (2.3) 也可写成点积的形式. 若令

$$\boldsymbol{N}=(b,-a),\quad \boldsymbol{X}=(x,y),\quad \boldsymbol{P}=(p,q)$$

那么方程 (2.3) 化为 $(\boldsymbol{X}-\boldsymbol{P})\cdot\boldsymbol{N}=0$, 即

$$\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{N}.$$

因为 $\boldsymbol{N}\cdot\boldsymbol{A}=ba-ab=0$, 所以向量 $\boldsymbol{N}$ 垂直于方向向量 $\boldsymbol{A}=(a,b)$, 我们称 $\boldsymbol{N}$ 为该直线的**法向量**(normal vector). 该直线由所有满足方程 $\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{N}$ 的点 $\boldsymbol{X}$ 组成.

这个关系的几何意义如图 2.2 所示, 点 $\boldsymbol{P}$ 和点 $\boldsymbol{X}$ 在该直线上, 法向量 $\boldsymbol{N}$ 与 $\boldsymbol{X}-\boldsymbol{P}$ 正交. 该图表明在该直线上的所有点 $\boldsymbol{X}$ 中, 当 $\boldsymbol{X}$ 为点 $\boldsymbol{P}$ 在 $\boldsymbol{N}$ 上的投影时, 其长度 $\|\boldsymbol{X}\|$ 最小. 下面我们给出这个结论的代数证明.

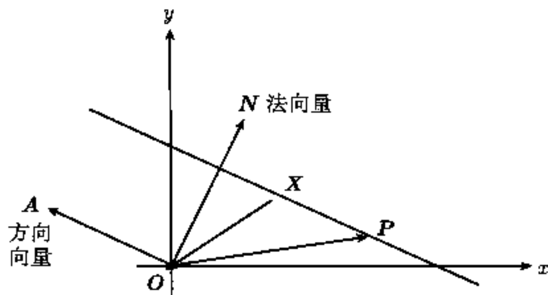


图 2.2 在 xy 平面中过点 $\boldsymbol{P}$, 法向量为 $\boldsymbol{N}$ 的直线, 该直线上的所有点满足 $\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{N}$

定理 2.6 令 L 为 $\mathbb{R}^2$ 中由满足以下关系的所有点 $\boldsymbol{X}$ 组成的直线:

$$\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{N},$$

其中 $\boldsymbol{P}$ 在该直线上且非零向量 $\boldsymbol{N}$ 为该直线的法向量. 令

$$d=\frac{|\boldsymbol{P}\cdot\boldsymbol{N}|}{\|\boldsymbol{N}\|},$$

则 L 上所有点 $\boldsymbol{X}$ 的长度 $\|\boldsymbol{X}\|\geqslant d$. 不仅如此, 当且仅当 $\boldsymbol{X}$ 是 $\boldsymbol{P}$ 在 $\boldsymbol{N}$ 上的投影时, 即当且仅当

$$\boldsymbol{X}=t\boldsymbol{N}$$

时, $\|\boldsymbol{X}\|=d$, 其中 $t=(\boldsymbol{P}\cdot\boldsymbol{N})/(\boldsymbol{N}\cdot\boldsymbol{N})$.

证明 因为 $\boldsymbol{X}\in L$, 所以 $\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{P}\cdot\boldsymbol{N}$, 由 Cauchy-Schwarz 不等式我们有

$$|\boldsymbol{P}\cdot\boldsymbol{N}|=|\boldsymbol{X}\cdot\boldsymbol{N}|\leqslant\|\boldsymbol{X}\|\,\|\boldsymbol{N}\|,$$

所以 $\|\boldsymbol{X}\|\geqslant|\boldsymbol{P}\cdot\boldsymbol{N}|/\|\boldsymbol{N}\|=d$. 又因为当且仅当存在纯量 t 使 $\boldsymbol{X}=t\boldsymbol{N}$ 时, 上式中等号成立, 此时 $\boldsymbol{P}\cdot\boldsymbol{N}=\boldsymbol{X}\cdot\boldsymbol{N}=t\boldsymbol{N}\cdot\boldsymbol{N}$, 所以 $t=\boldsymbol{P}\cdot\boldsymbol{N}/(\boldsymbol{N}\cdot\boldsymbol{N})$, 定理得证. □

类似地, 我们也可证明: 若 $\boldsymbol{Q}$ 为 $\mathbb{R}^2$ 中不在直线 L 上的点, 则对 L 上所有点 $\boldsymbol{X}$, $\|\boldsymbol{X}-\boldsymbol{Q}\|$ 的最小值为 $|(\boldsymbol{P}-\boldsymbol{Q})\cdot\boldsymbol{N}|/\|\boldsymbol{N}\|$, 此时 $\boldsymbol{X}-\boldsymbol{Q}$ 是 $\boldsymbol{P}-\boldsymbol{Q}$ 在法向量 $\boldsymbol{N}$ 上的投影. 数

$$\frac{|(\boldsymbol{P}-\boldsymbol{Q})\cdot\boldsymbol{N}|}{\|\boldsymbol{N}\|}$$

称为点 $\boldsymbol{Q}$ 到直线 L 的距离. 读者可用类似于图 2.2 的图说明这些概念.

2.6 习　题

1. 设 $\mathbb{R}^2$ 中的直线 L 包含点 $\boldsymbol{P}=(-3,1)$ 和点 $\boldsymbol{Q}=(1,1)$, 确定下列各点中哪些在 L 上:
   (a) (0,0)　(b) (0,1)
   (c) (1,2)　(d) (2,1)
   (e) $(-2,1)$.
2. 若 $\boldsymbol{P}=(2,-1)$, $\boldsymbol{Q}=(-4,2)$, 试确定习题 1 各点中哪些点在 L 上.
3. 设 $\mathbb{R}^3$ 中的直线 L 包含点 $\boldsymbol{P}=(-3,1,1)$ 且平行于向量 $(1,-2,3)$, 确定下列各点中哪些在 L 上:
   (a) (0,0,0)　(b) $(2,-1,4)$
   (c) $(-2,-1,4)$　(d) $(-4,3,-2)$
   (e) $(2,-9,16)$
4. 设 $\mathbb{R}^3$ 中的直线 L 包含点 $\boldsymbol{P}=(-3,1,1)$ 和 $\boldsymbol{Q}=(1,2,7)$, 确定下列各点中哪些在 L 上:
   (a) $(-7,0,5)$　(b) $(-7,0,-5)$
   (c) $(-11,1,11)$　(d) $(-11,-1,11)$
   (e) $(-1,\frac{3}{2},4)$　(f) $(-\frac{5}{3},\frac{4}{3},3)$
   (g) $(-1,\frac{3}{2},-4)$.
5. 判断下列各小题中三点 $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ 是否在 $\mathbb{R}^3$ 中的同一条直线上:
   (a) $\boldsymbol{P}=(2,1,1),\boldsymbol{Q}=(4,1,-1),\boldsymbol{R}=(3,-1,1)$.
   (b) $\boldsymbol{P}=(2,2,3),\boldsymbol{Q}=(-2,3,1),\boldsymbol{R}=(-6,4,1)$.
   (c) $\boldsymbol{P}=(2,1,1),\boldsymbol{Q}=(-2,3,1),\boldsymbol{R}=(5,-1,1)$.
6. 在 $\mathbb{R}^3$ 中的下列 8 个点中, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 在同一条直线上. 举出由这 8 个点组成的集合的所有至少包含三个点且其中所有点都在同一条直线上的子集: $\boldsymbol{A}=(2,1,1)$, $\boldsymbol{B}=(6,-1,1)$, $\boldsymbol{C}=(-6,5,1)$, $\boldsymbol{D}=(-2,3,1)$, $\boldsymbol{E}=(1,1,1)$, $\boldsymbol{F}=(-4,4,1)$, $\boldsymbol{G}=(-13,9,1)$, $\boldsymbol{H}=(14,-6,1)$.
7. 设 $\mathbb{R}^3$ 中的一条直线过点 $\boldsymbol{P}=(1,1,1)$ 且平行于向量 $\boldsymbol{A}=(1,2,3)$, 另一条直线过点 $\boldsymbol{Q}=(2,1,0)$ 且平行于向量 $\boldsymbol{B}=(3,8,13)$, 证明这两条直线相交并求它们的交点.
8. (a) 设 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{Q};\boldsymbol{B})$ 为 $\mathbb{R}^n$ 中的两条直线, 证明当且仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集包含 $\boldsymbol{P}-\boldsymbol{Q}$ 时, 这两条直线相交.
   (b) 判断 $\mathbb{R}^3$ 中如下两条直线是否相交:
   $$L=\{(1,1,-1)+t(-2,1,3)\},L'=\{(3,-4,1)+t(-1,5,2)\}.$$
9. 令 $\boldsymbol{X}(t)=\boldsymbol{P}+t\boldsymbol{A}$ 为直线 $L(\boldsymbol{P};\boldsymbol{A})$ 上的任意点, 其中 $\boldsymbol{P}=(1,2,3)$, $\boldsymbol{A}=(1,-2,2)$, 再令 $\boldsymbol{Q}=(3,3,1)$.
   (a) 计算 $\boldsymbol{Q}$ 与 $\boldsymbol{X}(t)$ 之间距离的平方 $\|\boldsymbol{Q}-\boldsymbol{X}(t)\|^2$.
   (b) 证明只有唯一的一点 $\boldsymbol{X}(t_0)$ 使距离 $\|\boldsymbol{Q}-\boldsymbol{X}(t)\|$ 取最小值, 并求此最小值.
   (c) 证明 $\boldsymbol{Q}-\boldsymbol{X}(t_0)$ 与 $\boldsymbol{A}$ 正交.
10. 设 $\boldsymbol{Q}$ 为 $\mathbb{R}^n$ 中不在直线 $L(\boldsymbol{P};\boldsymbol{A})$ 上的一点.
   (a) 令 $f(t)=\|\boldsymbol{Q}-\boldsymbol{X}(t)\|^2$, 其中 $\boldsymbol{X}(t)=\boldsymbol{P}+t\boldsymbol{A}$, 证明 $f(t)$ 是 t 的二次多项式且只有一个 t 使 $f(t)$ 取最小值, 设此时 t 的值为 t_0.
   (b) 证明 $\boldsymbol{Q}-\boldsymbol{X}(t_0)$ 与 $\boldsymbol{A}$ 正交.
11. 给定 $\mathbb{R}^n$ 中两条相互平行的直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{Q};\boldsymbol{A})$, 证明要么 $L(\boldsymbol{P};\boldsymbol{A})=L(\boldsymbol{Q};\boldsymbol{A})$, 要么交 $L(\boldsymbol{P};\boldsymbol{A})\cap L(\boldsymbol{Q};\boldsymbol{A})$ 为空集.
12. 给定 $\mathbb{R}^n$ 中两条相互不平行的直线 $L(\boldsymbol{P};\boldsymbol{A})$ 和 $L(\boldsymbol{Q};\boldsymbol{B})$, 证明它们要么没有交点, 要么仅有一个交点.

2.7 n 维欧氏空间中的平面

我们定义 n 维空间中的直线为形如 $\{\boldsymbol{P}+t\boldsymbol{A}\}$ 的点的集合, 也就是将非零向量 $\boldsymbol{A}$ 的线性生成集中的所有向量与点 $\boldsymbol{P}$ 相加后得到的集合. 用类似的方法我们也可定义平面, 唯一的区别是此时我们是将两个线性无关向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集中的所有向量与 $\boldsymbol{P}$ 相加. 为了保证 $\mathbb{R}^n$ 包含两个线性无关的向量, 我们假定 $n \geqslant 2$. 我们的应用主要集中在 $n=3$ 的情形.

定义 设 M 为 $\mathbb{R}^n$ 上的一个点集, 如果存在一点 $\boldsymbol{P}$ 和两个线性无关的向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 使得

$$M=\{p+s\boldsymbol{A}+t\boldsymbol{B}: s \text{ 和 } t \text{ 为实数}\},$$

则称 M 为一个**平面** (plane).

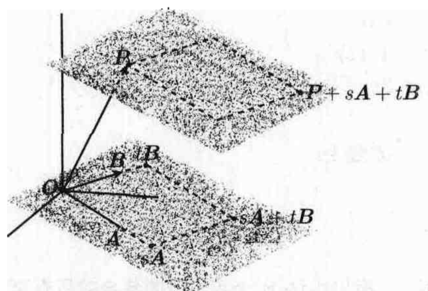


图 2.3 过 $\boldsymbol{P}$ 且由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面及其与过 $\boldsymbol{O}$ 且由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面的几何关系

为了简便起见, 我们用记号 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 表示这个平面, 称 M 中的各点在该平面上(on). 特别当我们令 $s=t=0$ 时, 我们知道 $\boldsymbol{P}$ 在该平面上. 我们也称点集 $\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 为过点 $\boldsymbol{P}$ 且由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面. 当 $\boldsymbol{P}=\boldsymbol{O}$ 时, 该平面就是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集. 图 2.3 画出了三维空间中一个由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成且过原点的平面以及一个由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成且过非零点 $\boldsymbol{P}$ 的平面.

接下来我们将推导平面的一些性质, 这些性质类似于定理 2.1~2.4 给出的直线的性质. 首先我们证明: 在定义平面 $\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 时, 可用线性生成集与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集相同的任意一对向量来代替 $\boldsymbol{A}$ 和 $\boldsymbol{B}$.

定理 2.7 设 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 和 $M'=\{\boldsymbol{P}+s\boldsymbol{C}+t\boldsymbol{D}\}$ 为过同一点 $\boldsymbol{P}$ 的两个平面, 那么当且仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集与 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集相同时, M 和 M' 相等.

证明 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集与 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集相同, 那么显然 $M=M'$. 要证逆命题, 设 $M=M'$, 因为平面 M 包含 $\boldsymbol{P}+\boldsymbol{A}$ 和 $\boldsymbol{P}+\boldsymbol{B}$, 而且这两点也在 M' 上, 所以 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 必在 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集中. 同理 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 也在 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集中, 所以 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集与 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集相同. □

下面的定理表明在定义平面 $\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 时, 可用该平面上的任意一点 $\boldsymbol{Q}$ 代替点 $\boldsymbol{P}$.

定理 2.8　*设 $M = \{\boldsymbol{P} + s\boldsymbol{A} + t\boldsymbol{B}\}$ 和 $M' = \{\boldsymbol{Q} + s\boldsymbol{A} + t\boldsymbol{B}\}$ 为由同一对向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的两个平面，那么当且仅当 $\boldsymbol{Q}$ 在 M 上时，这两个平面相同.*

证明　若 $M = M'$，则显然 $\boldsymbol{Q}$ 在 M 上. 要证逆命题，设 $\boldsymbol{Q}$ 在 M 上，令 $\boldsymbol{Q} = \boldsymbol{P} + a\boldsymbol{A} + b\boldsymbol{B}$. 取 M 上任意点 $\boldsymbol{X}$，则存在纯量 s 和 t，使 $\boldsymbol{X} = \boldsymbol{P} + s\boldsymbol{A} + t\boldsymbol{B}$. 又因为 $\boldsymbol{P} = \boldsymbol{Q} - a\boldsymbol{A} - b\boldsymbol{B}$，所以 $\boldsymbol{X} = \boldsymbol{Q} + (s-a)\boldsymbol{A} + (t-b)\boldsymbol{B}$，因此 X 在 M' 上，故 $M \subseteq M'$. 同理，我们有 $M' \subseteq M$，所以这两个平面相同. □

对平面也有类似于 Euclid 的平行公设 (定理 2.3) 的结果. 为了给出这个定理，我们必须先定义两个平面平行的概念. 图 2.2 向我们展示了两个平行平面的几何关系.

定义　*设 $M = \{\boldsymbol{P} + s\boldsymbol{A} + t\boldsymbol{B}\}$ 和 $M' = \{\boldsymbol{Q} + s\boldsymbol{C} + t\boldsymbol{D}\}$ 为两个平面，若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集与 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集相同，我们称这两个平面相互平行* (parallel). *若向量 $\boldsymbol{X}$ 在 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集中，我们也说 $\boldsymbol{X}$ 平行于平面 M.*

定理 2.9　*给定一个平面 M 和 M 外一点 $\boldsymbol{Q}$，有且仅有一个过 $\boldsymbol{Q}$ 且平行于 M 的平面 M'.*

证明　令 $M = \{\boldsymbol{P} + s\boldsymbol{A} + t\boldsymbol{B}\}$，考察平面 $M' = \{\boldsymbol{Q} + s\boldsymbol{A} + t\boldsymbol{B}\}$. 我们知道 M' 包含 $\boldsymbol{Q}$ 而且它由生成 M 的向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成，所以 M' 平行于 M. 再设 M'' 是过点 $\boldsymbol{Q}$ 且平行于 M 的另一个平面，那么

$$M'' = \{\boldsymbol{Q} + s\boldsymbol{C} + t\boldsymbol{D}\},$$

其中 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 的线性生成集与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集相同. 由定理 2.7，我们有 $M'' = M'$，所以 M' 是唯一一个过点 $\boldsymbol{Q}$ 且平行于 M 的平面. □

定理 2.4 告诉我们两个不同的点确定一条直线，下面的定理表明不在同一条直线上的三点确定一个平面.

定理 2.10　*设 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 为 n 维空间中不在同一直线上的三个点，那么有且仅有一个平面 M 包含这三个点. 可用如下集合表示 M:*

$$M = \{P + s(Q-P) + t(R-P)\}. \tag{2.4}$$

证明　首先我们假定三点中的某一点为原点 $\boldsymbol{O}$，不妨设 $\boldsymbol{P} = \boldsymbol{O}$，那么 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 不在过原点的同一条直线上，所以它们线性无关，因此它们生成一个过原点的平面，令这个平面为

$$M' = \{s\boldsymbol{Q} + t\boldsymbol{R}\},$$

显然这个平面包含 $\boldsymbol{O}, \boldsymbol{Q}, \boldsymbol{R}$ 三点.

现在我们来证明 M' 是包含 $\boldsymbol{O}, \boldsymbol{Q}, \boldsymbol{R}$ 三点的唯一一个平面. 所有其他过原点的平面 M'' 有如下形式:

$$M'' = \{s\boldsymbol{A} + t\boldsymbol{B}\},$$

其中 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性无关. 如果 M'' 包含 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$，则存在纯量 a, b, c, d，使

$$Q = aA + bB, \quad R = cA + dB, \tag{2.5}$$

所以 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的任意线性组合也是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性组合, 从而 $M' \subseteq M''$.

为了证明 $M'' \subseteq M'$, 我们只需证明 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的线性组合. 用 d 乘式 (2.5) 中第一个等式, 用 b 乘式 (2.5) 中第二个等式, 并将所得结果相减可消去 $\boldsymbol{B}$, 此时我们有

$$(ad-bc)\boldsymbol{A} = d\boldsymbol{Q} - b\boldsymbol{R}.$$

且 $ad-bc$ 不为零, 否则 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 线性相关, 所以我们可用 $(ad-bc)$ 除上式来将 $\boldsymbol{A}$ 表示为 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的线性组合. 类似地, 我们也可在式 (2.5) 中消去 $\boldsymbol{A}$ 而将 $\boldsymbol{B}$ 表示为 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$ 的线性组合, 所以 $M'' \subseteq M'$. 因此当 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 中有一个为原点时, 定理得证.

为了在一般情况下证明本定理, 我们令 M 为式 (2.4) 中的集合, 并令 $\boldsymbol{C} = \boldsymbol{Q} - \boldsymbol{P}$ 及 $\boldsymbol{D} = \boldsymbol{R} - \boldsymbol{P}$. 首先我们证明 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 线性无关. 否则存在纯量 t 使 $\boldsymbol{D} = t\boldsymbol{C}$, 所以 $\boldsymbol{R} - \boldsymbol{P} = t(\boldsymbol{Q} - \boldsymbol{P})$, 即 $\boldsymbol{R} = \boldsymbol{P} + t(\boldsymbol{Q} - \boldsymbol{P})$, 这与 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 不在同一条直线上矛盾. 所以 M 是过 $\boldsymbol{P}$ 且由线性无关向量 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 生成的平面, 这个平面包含 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 三点 (令 $s=1$ 与 $t=0$, 得 $\boldsymbol{Q}$ 在 M 上; 令 $s=0$ 与 $t=1$, 得 $\boldsymbol{R}$ 在 M 上). 现在我们必须证明它是唯一一个包含 $\boldsymbol{P}, \boldsymbol{Q}, \boldsymbol{R}$ 的平面.

令 M' 为包含 $\boldsymbol{P}, \boldsymbol{Q}$ 与 $\boldsymbol{R}$ 的任意平面, 由于 M' 包含 $\boldsymbol{P}$, 我们知道存在线性无关向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 使得

$$M' = \{\boldsymbol{P} + s\boldsymbol{A} + t\boldsymbol{B}\}.$$

令 $M_0 = \{s\boldsymbol{A} + t\boldsymbol{B}\}$ 为过原点且由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 这两个向量生成的平面. 显然当且仅当 M_0 包含 $\boldsymbol{X} - \boldsymbol{P}$ 时, M' 包含向量 $\boldsymbol{X}$. 因为 M' 包含 $\boldsymbol{Q}$ 和 $\boldsymbol{R}$, 所以 M_0 包含 $\boldsymbol{C} = \boldsymbol{Q} - \boldsymbol{P}$ 和 $\boldsymbol{D} = \boldsymbol{R} - \boldsymbol{P}$, 然而因为 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 线性无关, 我们已经证明了有且仅有一个包含 $\boldsymbol{O}, \boldsymbol{C}, \boldsymbol{D}$ 的平面, 所以 $M_0 = \{s\boldsymbol{C} + t\boldsymbol{D}\}$, 故 $M' = \{\boldsymbol{P} + s\boldsymbol{C} + t\boldsymbol{D}\} = M$, 定理得证. □

在定理 2.5 中, 我们证明了当且仅当 n 维空间中两个向量在同一条过原点的直线上时, 这两个向量线性相关. 下面的定理说明对三个向量有类似结论.

定理 2.11 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 为 n 维空间中的三个向量, 当且仅当它们在同一个过原点的平面上时, 它们线性相关.

证明 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 线性相关, 那么我们可将其中一个表示为另两个的线性组合, 不妨设 $\boldsymbol{C} = s\boldsymbol{A} + t\boldsymbol{B}$. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性无关, 那么它们生成一个过原点的平面, 且 $\boldsymbol{C}$ 在此平面上. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关, 那么 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 在同一条过原点的直线上, 所以它们在包含这条直线的任意平面上, 且该平面过原点.

为了证明逆命题, 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 在过原点的某一平面 M 上. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关, 则 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 线性相关, 定理得证. 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性无关, 则它们生成一个过原点的平面 M', 由定理 2.10, 有且仅有一个过原点且包含 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的平面, 所以 $M' = M$. 因为 $\boldsymbol{C}$ 在此平面上, 所以我们有 $\boldsymbol{C} = s\boldsymbol{A} + t\boldsymbol{B}$, 故 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 线性相关. □

2.8 平面和向量值函数

将一对实数 s 和 t 与平面 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 上向量 $\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}$ 联系起来的对应关系是向量值函数的又一个例子. 这里, 函数的定义域是 $\mathbb{R}^2$, 即所有实数对 (s,t) 的集合, 它的值域是平面 M. 若我们用 $\boldsymbol{X}$ 表示这个函数并用 $\boldsymbol{X}(s,t)$ 表示函数值, 那么对任意实数对 (s,t), 我们有

$$X(s,t)=P+sA+tB. \tag{2.6}$$

我们称 $\boldsymbol{X}$ 为两个实变量的向量值函数, 称纯量 s 和 t 为参数, 称式 (2.6) 为该平面的**参数方程**或**向量方程**. 与用一个实变量的向量值函数表示一条直线相同, 式 (2.6) 用两个参数的向量值函数表示一个平面.

如果我们用分量的形式表示向量方程 (2.6) 中的向量, 那么该向量方程相当于 n 个纯量方程. 例如, 在 $\mathbb{R}^3$ 中, 若令

$$\boldsymbol{P}=(p_1,p_2,p_3),\quad \boldsymbol{A}=(a_1,a_2,a_3),\quad \boldsymbol{B}=(b_1,b_2,b_3),\quad \boldsymbol{X}(s,t)=(x,y,z).$$

则向量方程 (2.6) 可以表示为如下三个纯量方程:

$$x=p_1+sa_1+tb_1,\quad y=p_2+sa_2+tb_2,\quad z=p_3+sa_3+tb_3.$$

我们总可以消去这些方程中的参数 s 和 t, 从而得到一个形如 $ax+by+cz=d$ 的线性方程, 这个方程称为**该平面的笛卡儿方程**, 如下例所示.

例 令 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$, 其中 $\boldsymbol{P}=(1,2,3)$, $\boldsymbol{A}=(1,2,1)$, $\boldsymbol{B}=(1,-4,-1)$, 那么相应的向量方程为

$$\boldsymbol{X}(s,t)=(1,2,3)+s(1,2,1)+t(1,-4,-1).$$

由此我们可得如下三个纯量参数方程:

$$x=1+s+t,\quad y=2+2s-4t,\quad z=3+s-t.$$

为了得到笛卡儿方程, 我们将第一个方程写为 $x-1=s+t$, 第三个方程写为 $z-3=s-t$, 将这两个方程相加得 $2s=x+z-4$, 再将它们相减, 则得 $2t=x-z+2$, 将这两个公式代入关于 y 的纯量参数方程中, 我们得到笛卡儿方程

$$x+y-3z=-6.$$

在 2.17 节中, 我们将对线性笛卡儿方程作进一步研究.

2.9 习 题

1. 给定 $\mathbb{R}^3$ 中平面 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$, 其中 $\boldsymbol{P}=(1,2,-3)$, $\boldsymbol{A}=(3,2,1)$, $\boldsymbol{B}=(1,0,4)$, 确定以下各点中哪些在 M 上.
   (a) (1,2,0)　　(b) (1,2,1)
   (c) (6,4,6)　　(d) (6,6,6)
   (e) $(6,6,-5)$.

2. 设 M 是 $\mathbb{R}^3$ 中由 $\boldsymbol{P}=(1,1,-1)$, $\boldsymbol{Q}=(3,3,2)$ 和 $\boldsymbol{R}=(3,-1,-2)$ 三点确定的平面, 确定以下各点中哪些在 M 上.
   (a) $(2,2,\frac{1}{2})$　　(b) $(4,0,-\frac{1}{2})$
   (c) $(-3,1-3)$　　(d) $(3,1,3)$
   (e) $(0,0,0)$.
3. 求 $\mathbb{R}^3$ 中下述各平面的纯量参数方程:
   (a) 过点 $(1,2,1)$ 由向量 $(0,1,0)$ 和 $(1,1,4)$ 生成的平面.
   (b) 过三点 $(1,2,1)$, $(0,1,0)$, $(1,1,4)$ 的平面.
4. 设 $\mathbb{R}^3$ 中平面 M 的纯量参数方程为
$$x=1+s-2t,\quad y=2+s+4t,\quad z=2s+t.$$
   (a) 确定以下哪些点在 M 上: $(0,0,0)$, $(1,2,0)$, $(2,-3,-3)$.
   (b) 求向量 $\boldsymbol{P}$, $\boldsymbol{A}$, $\boldsymbol{B}$, 使得 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$.
5. 令 M 为 $\mathbb{R}^3$ 中由不在同一直线上的三点 $\boldsymbol{P}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ 确定的平面,
   (a) 若 p, q, r 是和为 1 的三个纯量, 证明 $p\boldsymbol{P}+q\boldsymbol{Q}+r\boldsymbol{R}$ 在 M 上.
   (b) 证明 M 上所有点都可写为 $p\boldsymbol{P}+q\boldsymbol{Q}+r\boldsymbol{R}$ 的形式, 其中 $p+q+r=1$.
6. 求 $\mathbb{R}^3$ 中以下平面的形如 $ax+by+cz=d$ 的笛卡儿方程.
   (a) 过点 $(2,3,1)$ 由 $(3,2,1)$ 和 $(-1,-2,-3)$ 生成的平面.
   (b) 过点 $(2,3,1)$, $(-2,-1,-3)$ 与 $(4,3,-1)$ 的平面.
   (c) 过点 $(2,3,1)$ 且与过原点由 $(2,0,-2)$, $(1,1,1)$ 生成的平面平行的平面.
7. 设 $\mathbb{R}^3$ 中平面 M 的笛卡儿方程为 $3x-5y+z=9$.
   (a) 确定以下各点中哪些在 M 上: $(0,-2,-1)$, $(-1,-2,2)$, $(3,1,-5)$.
   (b) 求向量 $\boldsymbol{P}$, $\boldsymbol{A}$, $\boldsymbol{B}$, 使得 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$.
8. 给定 $\mathbb{R}^3$ 中两个平面 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$, $M'=\{\boldsymbol{Q}+s\boldsymbol{C}+t\boldsymbol{D}\}$, 其中 $\boldsymbol{P}=(1,1,1)$, $\boldsymbol{A}=(2,-1,3)$, $\boldsymbol{B}=(-1,0,2)$, $\boldsymbol{Q}=(2,3,1)$, $\boldsymbol{C}=(1,2,3)$, $\boldsymbol{D}=(3,2,1)$, 找出它们的交集 $M\cap M'$ 中的两个不同点.
9. 在 $\mathbb{R}^3$ 中给定平面 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$, 其中 $\boldsymbol{P}=(2,3,1)$, $\boldsymbol{A}=(1,2,3)$, $\boldsymbol{B}=(3,2,1)$, 以及笛卡儿方程为 $x-2y+z=0$ 的平面 M',
   (a) 确定 M 和 M' 是否平行.
   (b) 如果第三个平面 M'' 的笛卡儿方程为 $x+2y+z=0$, 找出 M' 和 M'' 的交集中的两个不同点.
10. 在 $\mathbb{R}^3$ 中, 令 L 为过点 $(1,1,1)$ 且平行于向量 $(2,-1,3)$ 的直线, 并令 M 为过点 $(1,1,-2)$ 由向量 $(2,1,3)$ 和 $(0,1,1)$ 生成的平面, 证明交集 $L\cap M$ 有且仅有一点并计算该点.
11. 在 $\mathbb{R}^3$ 中, 若向量 $\boldsymbol{X}$ 平行于平面 M, 那么称以 $\boldsymbol{X}$ 为方向向量的直线平行于 M. 令 L 为过点 $(1,1,1)$ 且平行于向量 $(2,-1,3)$ 的直线, 确定 L 是否平行于以下各平面.
    (a) 过点 $(1,1,-2)$ 由 $(2,1,3)$ 和 $(\frac{3}{4},1,1)$ 生成的平面.
    (b) 过点 $(1,1,-2)$, $(3,5,2)$, $(2,4,-1)$ 的平面.
    (c) 笛卡儿方程为 $x+2y+3z=-3$ 的平面.
12. 在 $\mathbb{R}^n$ 中, 设两个不同的点 $\boldsymbol{P}$ 和点 $\boldsymbol{Q}$ 在平面 M 中, 证明或推翻下述命题: 对于过 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 的直线, 其上的任意点都在 M 中.
13. 在 $\mathbb{R}^3$ 中, 给定过点 $(1,2,3)$ 且平行于向量 $(1,1,1)$ 的直线 L 和不在 L 上的点 $(2,3,5)$, 求过点 $(2,3,5)$ 且包含所有 L 上点的平面的笛卡儿方程.
14. 在 $\mathbb{R}^n$ 中, 给定直线 L 和不在 L 上的一点 $\boldsymbol{P}$, 证明或推翻下述命题: 有且仅有一个过 $\boldsymbol{P}$ 且包含 L 上所有点的平面.

2.10 $\mathbb{R}^3$ 中两向量的叉积

当用向量代数解决几何和力学上的问题时, 我们经常需要有一种简单方法来构造与给定的两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都垂直的向量. 我们用**叉积** (cross product) $\boldsymbol{A}\times\boldsymbol{B}$ (读作"$\boldsymbol{A}$ 叉 $\boldsymbol{B}$") 来解决这个问题, 叉积的定义如下.

定义 令 $\boldsymbol{A}=(a_1,a_2,a_3)$, $\boldsymbol{B}=(b_1,b_2,b_3)$ 为三维空间中的两个向量, 定义它们的**叉积** $\boldsymbol{A}\times\boldsymbol{B}$ (按此顺序) 为向量

$$A\times B=(a_2b_3-a_3b_2,a_3b_1-a_1b_3,a_1b_2-a_2b_1).$$

由定义可得叉积的下述性质.

定理 2.12 对三维空间中的任意向量 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 以及任意实数 c, 我们有

(a) $\boldsymbol{A}\times\boldsymbol{B}=-(\boldsymbol{B}\times\boldsymbol{A})$ (反对称性),

(b) $\boldsymbol{A}\times(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}\times\boldsymbol{B})+(\boldsymbol{A}\times\boldsymbol{C})$ (分配律),

(c) $c(\boldsymbol{A}\times\boldsymbol{B})=(c\boldsymbol{A})\times\boldsymbol{B}=\boldsymbol{A}\times(c\boldsymbol{B})$,

(d) $\boldsymbol{A}\cdot(\boldsymbol{A}\times\boldsymbol{B})=0$ (与 $\boldsymbol{A}$ 的正交性),

(e) $\boldsymbol{B}\cdot(\boldsymbol{A}\times\boldsymbol{B})=0$ (与 $\boldsymbol{B}$ 的正交性),

(f) $\|\boldsymbol{A}\times\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2-(\boldsymbol{A}\cdot\boldsymbol{B})^2$ (Lagrange 公式),

(g) 当且仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关时, $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{O}$.

证明 由定义易得性质 (a), (b), (c), 它们的证明留给读者作为练习. 现在我们来证性质 (d). 我们知道

$$\boldsymbol{A}\cdot(\boldsymbol{A}\times\boldsymbol{B})=a_1(a_2b_3-a_3b_2)+a_2(a_3b_1-a_1b_3)+a_3(a_1b_2-a_2b_1)=0.$$

同理可证性质 (e), 或者由性质 (a) 和 (d) 推出 (e). 为了证明性质 (f), 由定义我们有

$$\|\boldsymbol{A}\times\boldsymbol{B}\|^2=(a_2b_3-a_3b_2)^2+(a_3b_1-a_1b_3)^2+(a_1b_2-a_2b_1)^2$$

以及

$$\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2-(\boldsymbol{A}\cdot\boldsymbol{B})^2=(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2$$

将以上两式展开可知它们的右边相等.

性质 (f) 说明当且仅当 $(\boldsymbol{A}\cdot\boldsymbol{B})^2=\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2$ 时 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{O}$. 又由 Cauchy-Schwarz 不等式 (定理 1.3), $(\boldsymbol{A}\cdot\boldsymbol{B})^2=\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2$ 的充要条件是其中一个向量是另一个的纯量倍数, 即当且仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关时 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{O}$, 所以性质 (g) 得证. □

例 性质 (a) 和 (g) 都说明 $\boldsymbol{A}\times\boldsymbol{A}=\boldsymbol{O}$. 由叉积的定义, 我们知道单位坐标向量满足

$$\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k},\quad \boldsymbol{j}\times\boldsymbol{k}=\boldsymbol{i},\quad \boldsymbol{k}\times\boldsymbol{i}=\boldsymbol{j}.$$

叉积不满足结合律, 例如

$$i \times (i \times j) = i \times k = -j \quad 但 \quad (i \times i) \times j = O. \qquad \square$$

下面的定理给出了叉积的另外两条性质.

定理 2.13　令 A 和 B 为三维空间中的两个线性无关的向量. 则

(a) 向量 $A, B, A \times B$ 线性无关,

(b) 三维空间中任意一个与 A 和 B 都垂直的向量 N 都是 $A \times B$ 的纯量倍数.

证明　令 $C = A \times B$, 因为 A 和 B 线性无关, 所以 $C \neq O$. 设 $aA+bB+cC = O$, 用 C 点乘此式两边, 再由 $A \cdot C = B \cdot C = 0$ 可知 $c = 0$, 由此可得 $aA+bB = O$. 又因为 A 和 B 线性无关, 所以 $a = b = 0$, 因此 (a) 得证.

为了证明 (b), 设 N 与 A 和 B 都垂直, 并令 $C = A \times B$. 我们需要证明

$$(N \cdot C)^2 = (N \cdot N)(C \cdot C).$$

再由 Cauchy-Schwarz 不等式 (定理 1.3) 可知 N 是 C 的纯量倍数.

因为 A, B, C 线性无关, 由定理 1.10(c), 我们知道它们生成 $\mathbb{R}^3$, 所以存在纯量 a, b, c, 使得

$$N = aA + bB + cC.$$

又因为 $N \cdot A = N \cdot B = 0$, 所以

$$N \cdot N = N \cdot (aA + bB + cC) = cN \cdot C$$

同理, 因为 $C \cdot A = C \cdot B = 0$, 我们有

$$C \cdot N = C \cdot (aA + bB + cC) = cC \cdot C.$$

所以 $(N \cdot N)(C \cdot C) = (cN \cdot C)(C \cdot C) = (N \cdot C)(cC \cdot C) = (N \cdot C)^2$. 故定理得证. $\square$

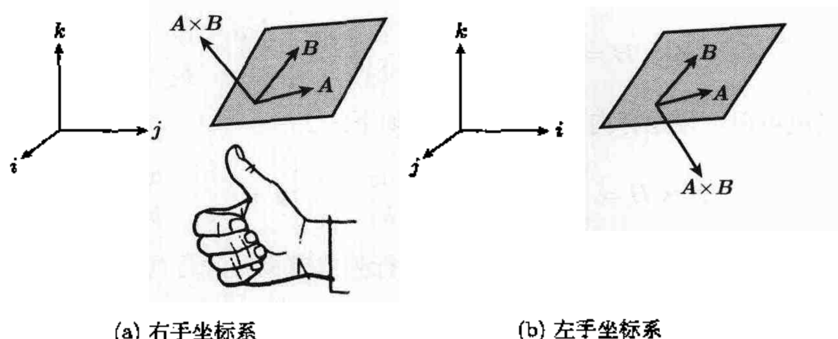
定理 2.12 给我们一个关于叉积的几何视角. 由性质 (d) 和 (e), 我们知道 $A \times B$ 与 A 和 B 都垂直. 若我们用一个箭头给出 $A \times B$ 的几何表示, 那么这个箭头的方向由三个坐标向量的相对位置确定. 如果 i, j, k 的顺序如图 2.4a 所示, 我们称它们构成了一个**右手坐标系** (right-handed coordinate system), 在这种情况下, $A \times B$ 的方向由"**右手法则** (right-hand rule)"确定. 即若用右手的拇指之外的手指表示 A 到 B 的旋转方向时, 则右手的拇指表示 $A \times B$ 的方向 (在此讨论中, 我们假定拇指与其余手指垂直). 在**左手坐标系**中, 如图 2.4b 所示, $A \times B$ 的方向与在右手坐标系中相反, 它由**左手法则**确定.

$A \times B$ 的长度有一个有趣的几何解释. 若 A 和 B 为非零向量且它们的夹角为 θ, 其中 $0 \leqslant \theta \leqslant \pi$, 我们将 $A \cdot B = \|A\|\|B\|\cos\theta$ 代入定理 2.12 中的性质 (f), 那么我们有

$$\|A \times B\|^2 = \|A\|^2\|B\|^2(1 - \cos^2\theta) = \|A\|^2\|B\|^2\sin^2\theta,$$

从而有

$$\|A \times B\| = \|A\|\,\|B\|\sin\theta.$$



(a) 右手坐标系 (b) 左手坐标系

图 2.4 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{A}\times\boldsymbol{B}$ 的相对位置

因为 $\|\boldsymbol{B}\|\sin\theta$ 是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 构成的平行四边形的高 (见图 2.5), 于是我们可知 $\boldsymbol{A}\times\boldsymbol{B}$ 的长度等于该平行四边形的面积.

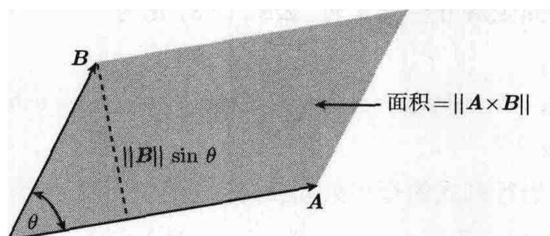


图 2.5 $\boldsymbol{A}\times\boldsymbol{B}$ 的长度等于由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 确定的平行四边形的面积

2.11 用行列式表示叉积

在行列式的帮助下, 叉积的定义公式可写成更紧凑的形式. 若 a, b, c, d 为四个数, 那么我们经常用记号

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

表示差 $ad-bc$, 我们称这个记号为一个 (二阶) **行列式** (determinant), 称数 a, b, c, d 为该行列式的**元素** (element) 或**分量** (entry), 该行列式有两行 a, b 和 c, d 以及两列 a, c 和 b, d. 注意到, 如果我们交换行列式的两行或两列, 所得行列式与原行列式的正负号相反. 例如, 因为 $ad-bc=-(bc-ad)$, 我们有

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} b & a \\ d & c \end{vmatrix}.$$

我们可用二阶行列式表示叉积的各个分量, 此时定义 $\boldsymbol{A}\times\boldsymbol{B}$ 的公式化为

$$\boldsymbol{A}\times\boldsymbol{B}=\left(\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix},\begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix},\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\right).$$

此公式也可用单位坐标向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 表示如下：

$$\boldsymbol{A}\times\boldsymbol{B}=\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\boldsymbol{i}+\begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}\boldsymbol{j}+\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\boldsymbol{k}. \tag{2.7}$$

三阶行列式有三行三列, 可用下式由二阶行列式定义三阶行列式：

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}-a_2\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}+a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}. \tag{2.8}$$

上式称为将行列式按第一行展开. 注意到上式右边与 a_1 相乘的行列式可由左边的行列式划去 a_1 所在的行和列得到. 我们可类似得到上式右边的其他行列式.

当我们用分量完全展开行列式时, 公式 (2.8) 化为

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_2 \end{vmatrix}=a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1.$$

这个公式告诉我们当行列式的行与列互换时, 它的值不变, 换句话说, 我们有

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

我们将在第 5 章讨论阶数大于三的行列式. 这里我们引入二阶和三阶行列式的目的仅在于使有关公式易于记忆.

当行列式的第一行为向量时, 行列式仍有意义. 例如, 用式 (2.8) 中的展开法则展开行列式

$$\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

我们发现结果与式 (2.7) 的右边相同, 这个结果使我们可将叉积 $\boldsymbol{A}\times\boldsymbol{B}$ 的定义写成如下的紧凑形式：

$$\boldsymbol{A}\times\boldsymbol{B}=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

例如, 为了计算 $\boldsymbol{A}=2\boldsymbol{i}-8\boldsymbol{j}+3\boldsymbol{k}$ 和 $\boldsymbol{B}=4\boldsymbol{j}+3\boldsymbol{k}$ 的叉积, 我们有

$$A\times B=\begin{vmatrix} i & j & k\\ 2 & -8 & 3\\ 0 & 4 & 3\end{vmatrix}=\begin{vmatrix} -8 & 3\\ 4 & 3\end{vmatrix}i-\begin{vmatrix} 2 & 3\\ 0 & 3\end{vmatrix}j+\begin{vmatrix} 2 & -8\\ 0 & 4\end{vmatrix}k=-36i-6j+8k.$$

2.12 习　题

1. 令 $\boldsymbol{A}=-\boldsymbol{i}+2\boldsymbol{k}$, $\boldsymbol{B}=2\boldsymbol{i}+\boldsymbol{j}-\boldsymbol{k}$, $\boldsymbol{C}=\boldsymbol{i}+2\boldsymbol{j}+2\boldsymbol{k}$, 用 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 表示下列各向量.
   (a) $\boldsymbol{A}\times\boldsymbol{B}$　(b) $\boldsymbol{B}\times\boldsymbol{C}$
   (c) $\boldsymbol{C}\times\boldsymbol{A}$　(d) $\boldsymbol{A}\times(\boldsymbol{C}\times\boldsymbol{A})$
   (e) $(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}$　(f) $\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C})$
   (g) $(\boldsymbol{A}\times\boldsymbol{C})\times\boldsymbol{B}$　(h) $(\boldsymbol{A}+\boldsymbol{B})\times(\boldsymbol{A}-\boldsymbol{C})$
   (i) $(\boldsymbol{A}\times\boldsymbol{B})\times(\boldsymbol{A}\times\boldsymbol{C})$
2. 在下列各小题中, 求 $\mathbb{R}^3$ 中的一个长度为 1 且与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都垂直的向量.
   (a) $\boldsymbol{A}=\boldsymbol{i}+\boldsymbol{j}+\boldsymbol{k}$, $\boldsymbol{B}=2\boldsymbol{i}+3\boldsymbol{j}-\boldsymbol{k}$.
   (b) $\boldsymbol{A}=2\boldsymbol{i}-3\boldsymbol{j}+4\boldsymbol{k}$, $\boldsymbol{B}=-\boldsymbol{i}+5\boldsymbol{j}+7\boldsymbol{k}$.
   (c) $\boldsymbol{A}=\boldsymbol{i}-2\boldsymbol{j}+3\boldsymbol{k}$, $\boldsymbol{B}=-3\boldsymbol{i}+2\boldsymbol{j}-\boldsymbol{k}$.
3. 在下列各小题中, 用叉积计算以 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 为顶点的三角形的面积:
   (a) $\boldsymbol{A}=(0,2,2)$, $\boldsymbol{B}=(2,0,-1)$, $\boldsymbol{C}=(3,4,0)$.
   (b) $\boldsymbol{A}=(-2,3,1)$, $\boldsymbol{B}=(1,-3,4)$, $\boldsymbol{C}=(1,2,1)$.
   (c) $\boldsymbol{A}=(0,0,0)$, $\boldsymbol{B}=(0,1,1)$, $\boldsymbol{C}=(1,0,1)$.
4. 设 $\boldsymbol{A}=2\boldsymbol{i}+5\boldsymbol{j}+3\boldsymbol{k}$, $\boldsymbol{B}=2\boldsymbol{i}+7\boldsymbol{j}+4\boldsymbol{k}$, $\boldsymbol{C}=3\boldsymbol{i}+3\boldsymbol{j}+6\boldsymbol{k}$, 用 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 表示叉积 $(\boldsymbol{A}-\boldsymbol{C})\times(\boldsymbol{B}-\boldsymbol{A})$.
5. 证明当且仅当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交时, $\|\boldsymbol{A}\times\boldsymbol{B}\|=\|\boldsymbol{A}\|\|\boldsymbol{B}\|$.
6. 给定 $\mathbb{R}^3$ 中两个线性无关的向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 令 $\boldsymbol{C}=(\boldsymbol{B}\times\boldsymbol{A})-\boldsymbol{B}$.
   (a) 证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}+\boldsymbol{C}$ 正交.
   (b) 证明 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 的夹角 θ 满足 $\frac{1}{2}\pi<\theta<\pi$.
   (c) 若 $\|\boldsymbol{B}\|=1$, $\|\boldsymbol{B}\times\boldsymbol{A}\|=2$, 计算 $\boldsymbol{C}$ 的长度.
7. 令 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为 $\mathbb{R}^3$ 中两个长度都为 1 的正交向量.
   (a) 证明 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}\times\boldsymbol{B}$ 构成 $\mathbb{R}^3$ 中的一个标准正交基.
   (b) 令 $\boldsymbol{C}=(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{A}$, 证明 $\|\boldsymbol{C}\|=1$.
   (c) 画一个简图表明 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}\times\boldsymbol{B}$ 之间的几何关系, 并用这个简图证明如下关系:
   $$(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{A}=\boldsymbol{B},\quad (\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{B}=-\boldsymbol{A}.$$
   (d) 用代数方法证明 (c) 中给出的关系.
8. (a) 证明下述命题并给出几何解释: 若 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{O}$ 且 $\boldsymbol{A}\cdot\boldsymbol{B}=0$, 则 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 中至少有一个为零.
   (b) 给定 $\boldsymbol{A}\neq\boldsymbol{O}$, 若 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{A}\times\boldsymbol{C}$ 且 $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{A}\cdot\boldsymbol{C}$, 证明 $\boldsymbol{B}=\boldsymbol{C}$.
9. 令 $\boldsymbol{A}=2\boldsymbol{i}-\boldsymbol{j}+2\boldsymbol{k}$, $\boldsymbol{C}=3\boldsymbol{i}+4\boldsymbol{j}-\boldsymbol{k}$.
   (a) 求向量 $\boldsymbol{B}$ 使得 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{C}$, 本题是否有多解?
   (b) 求向量 $\boldsymbol{B}$ 使得 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{C}$ 且 $\boldsymbol{A}\cdot\boldsymbol{B}=1$, 本题是否有多解?
10. 给定 $\mathbb{R}^3$ 中一个向量 $\boldsymbol{A}$ 和一个与 $\boldsymbol{A}$ 正交的向量 $\boldsymbol{C}$, 证明有且仅有一个向量 $\boldsymbol{B}$ 使得 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{C}$ 且 $\boldsymbol{A}\cdot\boldsymbol{B}=1$.
11. 设一个平行四边形的三个顶点为 $\boldsymbol{A}=(1,0,1)$, $\boldsymbol{B}=(-1,1,1)$, $\boldsymbol{C}=(2,-1,2)$.
    (a) 求所有可能作为此平行四边形第四个顶点的点 $\boldsymbol{D}$.
    (b) 计算三角形 $\boldsymbol{ABC}$ 的面积.
12. 给定 $\mathbb{R}^3$ 中两个互不平行的向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 且 $\boldsymbol{A}\cdot\boldsymbol{B}=2$, $\|\boldsymbol{A}\|=1$, $\|\boldsymbol{B}\|=4$, 令 $\boldsymbol{C}=2(\boldsymbol{A}\times\boldsymbol{B})-3\boldsymbol{B}$,

计算 $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})$, $\|\boldsymbol{C}\|$ 以及 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的夹角的余弦值.

13. 给定 $\mathbb{R}^3$ 中两个线性无关的向量, 判断下述各命题是否永真.
    (a) $\boldsymbol{A}+\boldsymbol{B}$, $\boldsymbol{A}-\boldsymbol{B}$, $\boldsymbol{A}\times\boldsymbol{B}$ 线性无关.
    (b) $\boldsymbol{A}+\boldsymbol{B}$, $\boldsymbol{A}+(\boldsymbol{A}\times\boldsymbol{B})$, $\boldsymbol{B}+(\boldsymbol{A}\times\boldsymbol{B})$ 线性无关.
    (c) $\boldsymbol{A}$, $\boldsymbol{B}$, $(\boldsymbol{A}+\boldsymbol{B})\times(\boldsymbol{A}-\boldsymbol{B})$ 线性无关.
14. (a) 设 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 为 $\mathbb{R}^3$ 中三个向量, 证明当且仅当 $(\boldsymbol{B}-\boldsymbol{A})\times(\boldsymbol{C}-\boldsymbol{A})=\boldsymbol{O}$ 时, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 在同一条直线上.
    (b) 若 $\boldsymbol{A}\neq\boldsymbol{B}$, 证明同时过 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的直线由所有满足 $(\boldsymbol{P}-\boldsymbol{A})\times(\boldsymbol{P}-\boldsymbol{B})=\boldsymbol{O}$ 的点 $\boldsymbol{P}$ 组成.
15. 给定 $\mathbb{R}^3$ 中两个长度都为 1 的正交向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 令 $\boldsymbol{P}$ 为满足等式 $\boldsymbol{P}\times\boldsymbol{B}=\boldsymbol{A}-\boldsymbol{P}$ 的任意向量, 证明下述各命题.
    (a) $\boldsymbol{P}$ 与 $\boldsymbol{B}$ 正交且长度为 $\frac{1}{2}\sqrt{2}$.
    (b) $\boldsymbol{P}$, $\boldsymbol{B}$, $\boldsymbol{P}\times\boldsymbol{B}$ 组成 $\mathbb{R}^3$ 的一组基.
    (c) $(\boldsymbol{P}\times\boldsymbol{B})\times\boldsymbol{B}=-\boldsymbol{P}$.
    (d) $\boldsymbol{P}=\frac{1}{2}\boldsymbol{A}-\frac{1}{2}(\boldsymbol{A}\times\boldsymbol{B})$.

2.13 纯量三重积

我们可将点积和叉积组合起来构成**纯量三重积** (scalar triple product) $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$. 注意到 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$ 只有一种有意义的形式: $\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})$, 所以它是两个向量的点积, 因此它的值是一个纯量. 我们可用行列式法求纯量三重积的值. 设 $\boldsymbol{A}=(a_1,a_2,a_3)$, $\boldsymbol{B}=(b_1,b_2,b_3)$, $\boldsymbol{C}=(c_1,c_2,c_3)$ 并按公式 (2.7) 展开 $\boldsymbol{B}\times\boldsymbol{C}$, 然后与 $\boldsymbol{A}$ 点乘, 我们得到

$$\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}+a_2\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix}+a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_c \\ c_1 & c_2 & c_3 \end{vmatrix}$$

所以纯量三重积 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$ 等于各行由 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的分量组成的行列式的值.

在定理 2.12 中, 我们知道 $\mathbb{R}^3$ 中两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性相关的充要条件是它们的叉积 $\boldsymbol{A}\times\boldsymbol{B}$ 为零向量. 下面的定理给出了判断 $\mathbb{R}^3$ 中三个向量是否线性相关的类似准则.

定理 2.14 设 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 为 $\mathbb{R}^3$ 中三个向量, 当且仅当

$$\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=0$$

时, 它们线性相关.

证明 首先我们假定 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 线性相关. 如果 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 线性相关, 那么 $\boldsymbol{B}\times\boldsymbol{C}=\boldsymbol{O}$, 所以 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=0$. 如果 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 线性无关, 则因为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 线性相关, 所以存在不全为零的纯量 a, b, c, 使得 $a\boldsymbol{A}+b\boldsymbol{B}+c\boldsymbol{C}=\boldsymbol{O}$. 此式中 a 必不为零, 否则我们有 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 线性相关. 所以我们可用 a 除此式两边从而将 $\boldsymbol{A}$ 表示为 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的线性组合. 设 $\boldsymbol{A}=t\boldsymbol{B}+s\boldsymbol{C}$, 用 $\boldsymbol{B}\times\boldsymbol{C}$ 点乘此式两边, 因为 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 都与 $\boldsymbol{B}\times\boldsymbol{C}$ 正

交, 所以我们有

$$\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})=t\boldsymbol{B}\cdot\boldsymbol{B}\times\boldsymbol{C}+s\boldsymbol{C}\cdot\boldsymbol{B}\times\boldsymbol{C}=0,$$

因此由 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 线性相关可推得纯量三重积 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$ 等于零.

接下来证明逆命题. 假定 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=0$, 如果 B 和 C 线性相关, 那么 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 线性相关, 定理得证. 所以我们假定 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 线性无关, 则由定理 2.13, 三个向量 $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{B}\times\boldsymbol{C}$ 线性无关. 所以它们可生成 $\boldsymbol{A}$, 所以存在纯量 a, b, c, 使得

$$\boldsymbol{A}=a\boldsymbol{B}+b\boldsymbol{C}+c(\boldsymbol{B}\times\boldsymbol{C}).$$

用 $\boldsymbol{B}\times\boldsymbol{C}$ 点乘上式两边, 并由 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=0$ 可得 $c=0$, 所以 $\boldsymbol{A}=a\boldsymbol{B}+b\boldsymbol{C}$, 这就证明了 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 线性相关. □

例 为了判断三个向量 $(2,3,-1)$, $(3,-7,5)$, $(1,-5,2)$ 是否线性相关, 我们用行列式法计算它们的纯量三重积:

$$\begin{vmatrix}2 & 3 & -1\\ 3 & -7 & 5\\ 1 & -5 & 2\end{vmatrix}=2(-14+25)-3(6-5)-1(-15+7)=27.$$

因为它们的纯量三重积不为零, 所以这三个向量线性无关. □

纯量三重积有一个有趣的几何解释, 图 2.6 给出了不在同一平面上的三个几何向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, 它们确定了一个平行六面体. 这个六面体的高为 $\|\boldsymbol{C}\|\cos\phi$, 其中 ϕ 为 $\boldsymbol{A}\times\boldsymbol{B}$ 与 $\boldsymbol{C}$ 的夹角. 在此图中, 因为 ϕ 在区间 $0\leqslant\phi<\frac{1}{2}\pi$ 内, 所以 $\cos\phi$ 为正. 这个平行六面体的底面是一个平行四边形, 它的面积为 $\|\boldsymbol{A}\times\boldsymbol{B}\|$, 将底面面积乘以高即得此平行六面体的体积为

$$\|\boldsymbol{A}\times\boldsymbol{B}\|(\|\boldsymbol{C}\|\cos\phi)=(\boldsymbol{A}\times\boldsymbol{B})\cdot\boldsymbol{C}.$$

也就是说, 纯量三重积 $\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C}$ 等于由 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 确定的平行六面体的体积. 当 ϕ 在区间 $\frac{1}{2}\pi<\phi\leqslant\pi$ 内时 $\cos\phi$ 为负, 因此 $\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C}$ 等于平行六面体的体积的相反数. 所以对任意 ϕ, 平行六面体的体积为纯量三重积 $\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C}$ 的绝对值. 如果 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 在过原点的同一平面上, 那么它们线性相关且它们的纯量三重积为零. 在这种情形下, 平行六面体退化为平行四边形从而它的体积为零.

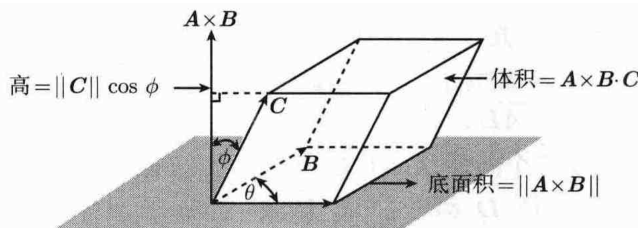


图 2.6 纯量三重积的几何解释: 平行六面体的体积

纯量三重积的几何解释也给我们关于纯量三重积的一些代数性质的启示. 例如, 对向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 作循环置换并不改变纯量三重积的值, 即

$$\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C}=\boldsymbol{B}\times\boldsymbol{C}\cdot\boldsymbol{A}=\boldsymbol{C}\times\boldsymbol{A}\cdot\boldsymbol{B} \tag{2.9}$$

我们将在 2.15 节的习题 8 中对这个性质的代数证明的要点给予提示. 该性质告诉我们纯量三重积的点乘号和叉乘号可互相交换. 事实上, 点积的交换律告诉我们 $(B\times C)\cdot A=\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})$, 再由式 (2.9) 中的第一个等号, 我们有

$$\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C}=\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}. \tag{2.10}$$

纯量三重积 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$ 常用记号 $[\boldsymbol{ABC}]$ 表示. 注意, 这个记号中并没有点乘号和叉乘号, 由等式 (2.10) 我们知道这个记号不会引起任何歧义; 因为纯量三重积只和三个乘数 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 的顺序有关而与点乘号和叉乘号的位置无关.

2.14 解三元线性方程组的 Cramer 法则

我们可用纯量三重积求解包含三个未知数 x, y, z 和三个联立的线性方程的方程组. 假定方程组具有如下形式:

$$\begin{aligned} a_1x+b_1y+c_1z&=d_1\\ a_2x+b_2y+c_2z&=d_2\\ a_3x+b_3y+c_3z&=d_3. \end{aligned} \tag{2.11}$$

令 $\boldsymbol{A}$ 为各分量分别为 a_1, a_2, a_3 的向量, 并类似地定义向量 $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, 那么 (2.11) 中的三个方程等价于向量方程

$$x\boldsymbol{A}+y\boldsymbol{B}+z\boldsymbol{C}=\boldsymbol{D}. \tag{2.12}$$

若我们用 $\boldsymbol{B}\times\boldsymbol{C}$ 点乘上式两边, 并用 $[\boldsymbol{ABC}]$ 表示 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$, 我们有

$$x[\boldsymbol{ABC}]+y[\boldsymbol{BBC}]+z[\boldsymbol{CBC}]=[\boldsymbol{DBC}].$$

又因为 $[\boldsymbol{BBC}]=[\boldsymbol{CBC}]=0$, 所以上式中包含 y 和 z 的各项的系数都为零, 我们可由此解出 x 为

$$x=\frac{[\boldsymbol{DBC}]}{[\boldsymbol{ABC}]},\quad [\boldsymbol{ABC}]\neq 0. \tag{2.13}$$

同理我们可得关于 y 和 z 的类似结果:

$$y=\frac{[\boldsymbol{ADC}]}{[\boldsymbol{ABC}]},\quad z=\frac{[\boldsymbol{ABD}]}{[\boldsymbol{ABC}]},\quad [\boldsymbol{ABC}]\neq 0. \tag{2.14}$$

条件 $[\boldsymbol{ABC}]\neq 0$ 表明三个向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 线性无关. 在这种情况下, 方程 (2.12) 表示三维空间中的任意向量 $\boldsymbol{D}$ 都可由 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 生成, 而且纯量 x, y 和 z 由公式 (2.13) 和 (2.14) 唯一确定. 当这些公式中的纯量三重积都写成行列式形式时, 结果即为求解方程组 (2.11) 的 Cramer 法则:

$$x=\frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}},\quad y=\frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}},\quad z=\frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

若 $[\boldsymbol{ABC}]=0$, 则 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 在过原点的同一平面上, 如果 $\boldsymbol{D}$ 不在该平面上, 则方程组无解. 若 $\boldsymbol{D}$ 与 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 在同一平面上, 易知方程组有无穷多解. 事实上, 向量 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 线性相关, 所以存在不全为零的纯量 u, v, w, 使得 $u\boldsymbol{A}+v\boldsymbol{B}+w\boldsymbol{C}=\boldsymbol{O}$. 如果三元组 (x,y,z) 满足方程 (2.12), 那么对任意实数 t, 我们有

$(x+tu)\boldsymbol{A}+(y+tv)\boldsymbol{B}+(z+tw)\boldsymbol{C}=x\boldsymbol{A}+y\boldsymbol{B}+z\boldsymbol{C}+t(u\boldsymbol{A}+v\boldsymbol{B}+w\boldsymbol{C})=x\boldsymbol{A}+y\boldsymbol{B}+z\boldsymbol{C}$.

所以三元组 $(x+tu, y+tv, z+tw)$ 也满足方程 (2.12).

2.15 习　题

1. 计算下列各小题的纯量三重积 $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}$.
   (a) $\boldsymbol{A}=(3,0,0)$, $\boldsymbol{B}=(0,4,0)$, $\boldsymbol{C}=(0,0,8)$.
   (b) $\boldsymbol{A}=(2,3,-1)$, $\boldsymbol{B}=(3,-7,5)$, $\boldsymbol{C}=(1,-5,2)$.
   (c) $\boldsymbol{A}=(2,1,3)$, $\boldsymbol{B}=(-3,0,6)$, $\boldsymbol{C}=(4,5,-1)$.
2. 求使向量 $(1,t,1)$, $(t,1,0)$, $(0,1,t)$ 线性相关的所有实数 t.
3. 计算由向量 $\boldsymbol{i}+\boldsymbol{j}$, $\boldsymbol{j}+\boldsymbol{k}$, $\boldsymbol{k}+\boldsymbol{i}$ 确定的平行六面体的体积.
4. 证明 $\boldsymbol{A}\times\boldsymbol{B}=\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{i})\boldsymbol{i}+\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{j})\boldsymbol{j}+\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{k})\boldsymbol{k}$.
5. 证明 $\boldsymbol{i}\times(\boldsymbol{A}\times\boldsymbol{i})+\boldsymbol{j}\times(\boldsymbol{A}\times\boldsymbol{j})+\boldsymbol{k}\times(\boldsymbol{A}\times\boldsymbol{k})=2\boldsymbol{A}$.
6. (a) 求满足关系
$$(a\boldsymbol{i}+b\boldsymbol{j}+c\boldsymbol{k})\cdot\boldsymbol{k}\times(6\boldsymbol{i}+3\boldsymbol{j}+4\boldsymbol{k})=3.$$
   的所有向量 $a\boldsymbol{i}+b\boldsymbol{j}+c\boldsymbol{k}$.
   (b) 求满足 (a) 中给出的关系的所有向量中长度最短的向量 $a\boldsymbol{i}+b\boldsymbol{j}+c\boldsymbol{k}$.
7. 用点积和叉积的代数性质推导纯量三重积的下述性质:
   (a) $(\boldsymbol{A}+\boldsymbol{B})\cdot(\boldsymbol{A}+\boldsymbol{B})\times\boldsymbol{C}=0$.
   (b) $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=-\boldsymbol{B}\cdot\boldsymbol{A}\times\boldsymbol{C}$(交换纯量三重积的前两个向量将改变结果的正负号).
   [提示: 用 (a) 的结果及分配律.]
   (c) $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=-\boldsymbol{A}\cdot\boldsymbol{C}\times\boldsymbol{B}$ (交换纯量三重积的后两个向量将改变结果的正负号).
   [提示: 用反对称性.]
   (d) $\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=-\boldsymbol{C}\cdot\boldsymbol{B}\times\boldsymbol{A}$ (交换纯量三重积的第一个和第三个向量将改变结果的正负号).
8. 用习题 7 中的性质 (b), (c), (d) 推导下述等式:
$$\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C}=\boldsymbol{B}\cdot\boldsymbol{C}\times\boldsymbol{A}=\boldsymbol{C}\cdot\boldsymbol{A}\times\boldsymbol{B}.$$
   这个结果表明 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 的循环置换不改变纯量三重积的值.
9. 本题勾画出向量恒等式
$$\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C})=(\boldsymbol{C}\cdot\boldsymbol{A})\boldsymbol{B}-(\boldsymbol{B}\cdot\boldsymbol{A})\boldsymbol{C}, \tag{2.15}$$

的证明要点, 这个公式有时称为"cab 减 bac"公式. 令 $\boldsymbol{B}=(b_1,b_2,b_3)$, $\boldsymbol{C}=(c_1,c_2,c_3)$, 证明:

$$\boldsymbol{i}\times(\boldsymbol{B}\times\boldsymbol{C})=c_1\boldsymbol{B}-b_1\boldsymbol{C}.$$

上式为当 $\boldsymbol{A}=\boldsymbol{i}$ 时恒等式 (2.15) 的特殊情形. 再证明当 $\boldsymbol{A}=\boldsymbol{j}$ 和 $\boldsymbol{A}=\boldsymbol{k}$ 时对应的特殊情形, 并将它们合并起来证明恒等式 (2.15).

10. 用公式 (2.15) 推导下述各向量恒等式:
    (a) $(\boldsymbol{A}\times\boldsymbol{B})\times(\boldsymbol{C}\times\boldsymbol{D})=(\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{D})\boldsymbol{C}-(\boldsymbol{A}\times\boldsymbol{B}\cdot\boldsymbol{C})\boldsymbol{D}$.
    (b) $\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C})+\boldsymbol{B}\times(\boldsymbol{C}\times\boldsymbol{A})+\boldsymbol{C}\times(\boldsymbol{A}\times\boldsymbol{B})=\boldsymbol{O}$.
    (c) $\boldsymbol{A}\times(\boldsymbol{B}\times\boldsymbol{C})=(\boldsymbol{A}\times\boldsymbol{B})\times\boldsymbol{C}$ 当且仅当 $\boldsymbol{B}\times(\boldsymbol{C}\times\boldsymbol{A})=\boldsymbol{O}$.
    (d) $(\boldsymbol{A}\times\boldsymbol{B})\cdot(\boldsymbol{C}\times\boldsymbol{D})=(\boldsymbol{B}\cdot\boldsymbol{D})(\boldsymbol{A}\cdot\boldsymbol{C})-(\boldsymbol{B}\cdot\boldsymbol{C})(\boldsymbol{A}\cdot\boldsymbol{D})$.
11. 设 $\mathbb{R}^3$ 中四个向量 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 满足关系 $\boldsymbol{A}\times\boldsymbol{C}\cdot\boldsymbol{B}=5$, $\boldsymbol{A}\times\boldsymbol{D}\cdot\boldsymbol{B}=3$, $\boldsymbol{C}+\boldsymbol{D}=\boldsymbol{i}+2\boldsymbol{j}+\boldsymbol{k}$, $\boldsymbol{C}-\boldsymbol{D}=\boldsymbol{i}-\boldsymbol{k}$, 用 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 表示 $(\boldsymbol{A}\times\boldsymbol{B})\times(\boldsymbol{C}\times\boldsymbol{D})$.
12. 证明 $(\boldsymbol{A}\times\boldsymbol{B})\cdot(\boldsymbol{B}\times\boldsymbol{C})\times(\boldsymbol{C}\times\boldsymbol{A})=(\boldsymbol{A}\cdot\boldsymbol{B}\times\boldsymbol{C})^2$.
13. 证明或推翻公式 $\boldsymbol{A}\times[\boldsymbol{A}\times(\boldsymbol{A}\times\boldsymbol{B})]\cdot\boldsymbol{C}=-\|\boldsymbol{A}\|^2\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})$.
14. (a) 证明顶点为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 的四面体的体积为

$$\frac{1}{6}|(\boldsymbol{B}-\boldsymbol{A})\cdot(\boldsymbol{C}-\boldsymbol{A})\times(\boldsymbol{D}-\boldsymbol{A})|.$$

    (b) 当 $\boldsymbol{A}=(1,1,1)$, $\boldsymbol{B}=(0,0,2)$, $\boldsymbol{C}=(0,3,0)$, $\boldsymbol{D}=(4,0,0)$ 时, 计算该四面体的体积.
15. (a) 若 $\boldsymbol{B}\neq\boldsymbol{C}$, 证明 $\boldsymbol{A}$ 到过 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 的直线的距离为

$$\|(\boldsymbol{A}-\boldsymbol{B})\times(\boldsymbol{C}-\boldsymbol{B})\|/\|\boldsymbol{B}-\boldsymbol{C}\|.$$

    (b) 当 $\boldsymbol{A}=(1,-2,-5)$, $\boldsymbol{B}=(-1,1,1)$, $\boldsymbol{C}=(4,5,1)$ 时, 计算这个距离的值.
16. Alexandia 的 Hero 给出了一个计算三条边长为 a, b, c 的三角形的面积 S 的公式 $S=\sqrt{s(s-a)(s-b)(s-c)}$, 其中 $s=(a+b+c)/2$. 本题给出了一个用向量方法证明该公式的方法. 假定三角形的各顶点为 $\boldsymbol{O}$, $\boldsymbol{A}$, $\boldsymbol{B}$, 其中 $\|\boldsymbol{A}\|=a$, $\|\boldsymbol{B}\|=b$, $\|\boldsymbol{B}-\boldsymbol{A}\|=c$,
    (a) 用恒等式 $\|\boldsymbol{A}\times\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2\|\boldsymbol{B}\|^2-(\boldsymbol{A}\cdot\boldsymbol{B})^2$ 和 $-2\boldsymbol{A}\cdot\boldsymbol{B}=\|\boldsymbol{A}-\boldsymbol{B}\|^2-\|\boldsymbol{A}\|^2-\|\boldsymbol{B}\|^2$ 推得公式

$$4S^2=a^2b^2-\frac{1}{4}(c^2-a^2-b^2)^2=\frac{1}{4}(2ab-c^2+a^2+b^2)(2ab+c^2-a^2-b^2).$$

    (b) 整理 (a) 中得到的公式从而获得公式

$$S^2=\frac{1}{16}(a+b+c)(a+b-c)(c-a+b)(c+a-b),$$

    并由此证明 Hero 公式.

用 Cramer 法则解习题 17~19 中的方程组

17. $x+2y+3z=5$, $2x-y+4z=11$, $-y+z=3$.
18. $x+y+2z=4$, $3x-y-z=2$, $2x+5y+3z=3$.
19. $x+y=5$, $x+z=2$, $y+z=5$.
20. 若 $\boldsymbol{P}=(1,1,1)$, $\boldsymbol{A}=(2,1,-1)$, 证明直线 $\{\boldsymbol{P}+t\boldsymbol{A}\}$ 上的各点 (x,y,z) 都满足线性方程组 $x-y+z=1$, $x+y+3z=5$, $3x+y+7z=11$.

2.16 $\mathbb{R}^3$ 中平面的法向量

在 2.6 节中, 我们将 n 维空间中的平面定义为形如 $\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 的集合, 其中 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是线性无关的向量. 接下来我们将看到可以用一种和前面所述完全不同的方法来定义三维空间中的平面, 这种方法将用到法向量的概念.

定义 令 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 为过点 $\boldsymbol{P}$ 由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面, $\boldsymbol{N}$ 为 $\mathbb{R}^3$ 中的向量, 若 $\boldsymbol{N}$ 与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都垂直, 则我们称 $\boldsymbol{N}$ **垂直于** M. 更进一步, 若 N 为非零向量, 则我们称 $\boldsymbol{N}$ 为 M 的**一个法向量** (normal vector).

注 若 $\boldsymbol{N}\cdot\boldsymbol{A}=\boldsymbol{N}\cdot\boldsymbol{B}=0$, 则 $\boldsymbol{N}\cdot(s\boldsymbol{A}+t\boldsymbol{B})=0$, 所以与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都垂直的向量与由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面上的任意向量都垂直. 此外, 若 $\boldsymbol{N}$ 为一个平面的法向量, 则对任意 $t\neq 0$, $t\boldsymbol{N}$ 也是该平面的法向量.

定理 2.15 在 $\mathbb{R}^3$ 中给定 $M=\{\boldsymbol{P}+s\boldsymbol{A}+t\boldsymbol{B}\}$ 为过点 $\boldsymbol{P}$ 由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面, 令 $\boldsymbol{N}=\boldsymbol{A}\times\boldsymbol{B}$, 那么我们有

(a) $\boldsymbol{N}$ 是 M 的一个法向量,

(b) M 为 $\mathbb{R}^3$ 中满足方程

$$(\boldsymbol{X}-\boldsymbol{P})\cdot\boldsymbol{N}=0 \tag{2.16}$$

的所有向量 $\boldsymbol{X}$ 组成的集合.

证明 因为 M 是一个平面, 所以 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 线性无关, 因此 $\boldsymbol{A}\times\boldsymbol{B}\neq\boldsymbol{O}$. 又因为 $\boldsymbol{A}\times\boldsymbol{B}$ 与 $\boldsymbol{A}$ 都正交, 所以 (a) 得证.

现在我们证明 (b). 令 M' 为 $\mathbb{R}^3$ 中满足方程 (2.16) 的所有向量组成的集合. 若 $\boldsymbol{X}\in M$, 则 $\boldsymbol{X}-\boldsymbol{P}$ 在 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性生成集中, 所以 $\boldsymbol{X}-\boldsymbol{P}$ 与 N 正交, 故有 $\boldsymbol{X}\in M'$, 于是 $M\subseteq M'$. 反之, 假定 $\boldsymbol{X}\in M'$, 则 X 满足方程 (2.16). 因为 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{N}$ 线性无关 (定理 2.13), 所以它们生成 $\mathbb{R}^3$ 中所有向量, 从而存在纯量 s, t, u, 使得

$$\boldsymbol{X}-\boldsymbol{P}=s\boldsymbol{A}+t\boldsymbol{B}+u\boldsymbol{N}.$$

用 $\boldsymbol{N}$ 点乘上式两边, 我们得 $u=0$, 所以 $\boldsymbol{X}-\boldsymbol{P}=s\boldsymbol{A}+t\boldsymbol{B}$. 此式表明 $\boldsymbol{X}\in M$, 所以 $M'\subseteq M$, 故定理得证. □

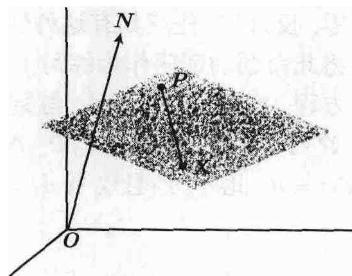


图 2.7 过点 $\boldsymbol{P}$ 和点 $\boldsymbol{X}$ 且以 $\boldsymbol{N}$ 为法向量的平面

定理 2.15 的几何意义如图 2.7 所示. 点 $\boldsymbol{P}$ 和点 $\boldsymbol{X}$ 在平面 M 上, 法向量 $\boldsymbol{N}$ 与 $\boldsymbol{X}-\boldsymbol{P}$ 正交. 该图也表明了下述定理.

定理 2.16 给定过点 $\boldsymbol{P}$ 的一个平面 M 和 M 的一个法向量 $\boldsymbol{N}$, 令

$$d=\frac{|\boldsymbol{P}\cdot\boldsymbol{N}|}{\|\boldsymbol{N}\|}. \tag{2.17}$$

则 M 上的任意点 $\boldsymbol{X}$ 的长度 $\|\boldsymbol{X}\|\geqslant d$. 更进一步我们有, 当且仅当 $\boldsymbol{X}$ 是 $\boldsymbol{P}$ 在 $\boldsymbol{N}$ 上的投影时, 即当且仅当

$$\boldsymbol{X}=t\boldsymbol{N},\quad 其中\quad t=\frac{\boldsymbol{P}\cdot\boldsymbol{N}}{\boldsymbol{N}\cdot\boldsymbol{N}}.$$

时, $\|\boldsymbol{X}\|=d$.

证明 定理 2.6 给出了关于 $\mathbb{R}^2$ 中直线的相应结果, 本定理可以用证明定理 2.6 的方法由 Cauchy-Schwarz 不等式导出. □

同理我们可知若 $\boldsymbol{Q}$ 为不在 M 上的点, 则对 M 上所有点 $\boldsymbol{X}$, 当 $\boldsymbol{X}-\boldsymbol{Q}$ 为 $\boldsymbol{P}-\boldsymbol{Q}$ 在 $\boldsymbol{N}$ 上的投影时, $\|\boldsymbol{X}-\boldsymbol{Q}\|$ 取最小值. 此最小值为 $|(\boldsymbol{P}-\boldsymbol{Q})\cdot\boldsymbol{N}|/\|\boldsymbol{N}\|$, 我们称它为 $\boldsymbol{Q}$ 到平面 M 的**距离**. 于是公式 (2.17) 中的数 d 即为原点到平面 M 的距离.

2.17 $\mathbb{R}^3$ 中平面的线性笛卡儿方程

我们也可将定理 2.15 和定理 2.16 的结果写成分量形式. 若 $\boldsymbol{N}=(a,b,c)$, $\boldsymbol{P}=(x_1,y_1,z_1)$, $\boldsymbol{X}=(x,y,z)$, 则向量方程 (2.16) 化为

$$a(x-x_1)+b(y-y_1)+c(z-z_1)=0. \tag{2.18}$$

上式称为**平面的笛卡儿方程**, 平面 M 上的点都满足此方程, 而且只有平面 M 上的点才满足此方程. 当我们用非零纯量 t 乘方程 (2.18) 两边后, 满足此式的点集并不改变, 其实这不过是在方程 (2.16) 中选取了不同的法向量 N.

我们令 $d_1=ax_1+by_1+cz_1$, 则方程 (2.18) 化为

$$ax+by+cz=d_1, \tag{2.19}$$

的形式, 称这样的方程为关于 x, y, z 的线性方程. 我们刚才已证明了三维空间中一平面上的任意点 (x,y,z) 满足一个形如 (2.19) 的线性笛卡儿方程, 其中, a, b, c 不全为零. 反过来, 任意具有这种性质的线性方程代表了三维空间中的一个平面. (读者可将此命题的证明作为练习.)

方程 (2.19) 中的数 d_1 与原点到该平面的距离 d 之间有一个简单关系. 由于 $d_1=\boldsymbol{P}\cdot\boldsymbol{N}$, 我们有 $|d_1|=|\boldsymbol{P}\cdot\boldsymbol{N}|=d\|\boldsymbol{N}\|$. 我们特别指出, 当向量 N 的长度为 1 时, $|d_1|=d$. 此外, 当且仅当 $d_1=0$ 时, 该平面过原点.

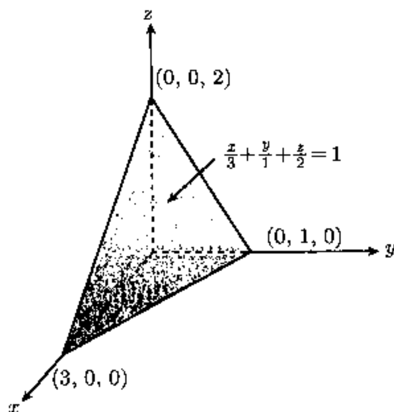


图 2.8 截距分别为 3, 1, 2 的平面

例 笛卡儿方程 $2x+6y+3z=6$ 代表一个以 $N=2\boldsymbol{i}+6\boldsymbol{j}+3\boldsymbol{k}$ 为法向量的平面. $\boldsymbol{N}$ 的长度为 $(\boldsymbol{N}\cdot\boldsymbol{N})^{1/2}=7$. 若我们用 6 除这个笛卡儿方程两边并将它化为

$$\frac{x}{3}+\frac{y}{1}+\frac{z}{2}=1,$$

我们看到该平面与各坐标轴分别交于 $(3,0,0)$, $(0,1,0)$, $(0,0,2)$. 数 3, 1, 2 分别称作该平面在 x, y, z 轴上的**截距**(intercept). 若知道一个平面的各截距, 我们可以快速地画出该平面的简图. 图 2.8 给出了该平面的一部分, 原点到该平面的距离 $d=6/\|\boldsymbol{N}\|=6/7$.

若两个平面相互平行, 则它们有公共法向量 $\boldsymbol{N}=(a,b,c)$, 且它们的笛卡儿方程可写作如下形式:

$$ax + by + cz = d_1, \quad ax + by + cz = d_2,$$

这两个笛卡儿方程唯一的区别仅在于等号右边的两个数. 受定理 2.16 的启发, 我们把数 $|d_1 - d_2|/\|\boldsymbol{N}\|$ 称为这两个平面的垂直距离.

当一个平面的法向量垂直于另一个平面的法向量时, 我们称这两个平面相互垂直. 更进一步, 若两个平面的法向量的夹角为 θ, 我们称这两个平面的夹角为 θ.

2.18 习 题

1. 给定向量 $\boldsymbol{A} = 2\boldsymbol{i} + 3\boldsymbol{j} - 4\boldsymbol{k}$ 和 $\boldsymbol{B} = \boldsymbol{j} + \boldsymbol{k}$,
   (a) 求一个与 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都垂直的非零向量 $\boldsymbol{N}$.
   (b) 求过原点由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 生成的平面的笛卡儿方程.
   (c) 求过点 $(1,2,3)$ 由 A 和 $\boldsymbol{B}$ 生成的平面的笛卡儿方程.
2. 已知一个平面的笛卡儿方程为 $x + 2y - 2z + 7 = 0$, 求
   (a) 该平面的一个长度为 1 的法向量, (b) 该平面的各截距,
   (c) 原点到该平面的距离, (d) 该平面上离原点最近的点 $\boldsymbol{Q}$.
3. 求过点 $(1,2,-3)$ 且与以 $3x - y + 2z = 4$ 为笛卡儿方程的平面平行的平面的笛卡儿方程. 计算这两个平面的垂直距离.
4. 设四个平面分别以 $x + 2y - 2z = 5$, $3x - 6y + 3z = 2$, $2x + y + 2z = -1$, $x - 2y + z = 7$ 为笛卡儿方程.
   (a) 证明它们之中有两个平面相互平行, 另外两个平面相互垂直.
   (b) 计算两个相互平行的平面之间的垂直距离.
5. 已知点 $(1,1,-1)$, $(3,3,2)$, $(3,-1,-2)$ 确定一个平面, 对于这个平面, 求
   (a) 它的一个法向量, (b) 它的一个笛卡儿方程,
   (c) 原点到它的距离.
6. 求由 $(1,2,3)$, $(2,3,4)$, $(-1,7,-2)$ 确定的平面的笛卡儿方程.
7. 计算分别以 $x + y = 1$ 和 $y + z = 2$ 为笛卡儿方程的两个平面的夹角.
8. 如果一个非零向量 $\boldsymbol{N}$ 是一个平面 M 的法向量, 那么我们说平行于 $\boldsymbol{N}$ 的所有向量都垂直于 M. 求过点 $(2,3,-7)$ 且垂直于过点 $(1,2,3)$ 和 $(2,4,12)$ 的直线的平面的笛卡儿方程.
9. 求过点 $(2,1,-3)$ 且垂直于由方程 $4x - 3y + z = 5$ 给出的平面的直线的向量参数方程.
10. 设有一个点在空间中运动, 它在时刻 t 的位置由向量 $\boldsymbol{X}(t) = (1-t)\boldsymbol{i} + (2-3t)\boldsymbol{j} + (2t-1)\boldsymbol{k}$ 给出.
    (a) 证明这个点沿着一条直线运动. (称这条直线为 L.)
    (b) 求一个平行于 L 的向量 $\boldsymbol{N}$.
    (c) 什么时候该点击中由方程 $2x + 3y + 2z + 1 = 0$ 给出的平面?
    (d) 求与 (c) 中给出的平面平行且包含点 $\boldsymbol{X}(3)$ 的平面的笛卡儿方程.
11. 若一个平面过点 $(1,1,1)$ 且它的法向量 $\boldsymbol{N}$ 与 $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ 的夹角分别为 $\frac{1}{3}\pi$, $\frac{1}{4}\pi$, $\frac{1}{3}\pi$, 求该平面的笛卡儿方程.
12. 计算以原点为一个顶点以由笛卡儿方程 $x + 2y + 3z = 6$ 给出的平面与各坐标轴的交点为另外三个顶点的四面体的体积.
13. 求长度为 1 的一个向量, 它垂直于 $\boldsymbol{i} + 2\boldsymbol{j} - 3\boldsymbol{k}$ 且平行于以 $x - y + 5z = 1$ 为笛卡儿方程的平面.
14. 求与向量 $\boldsymbol{i} + \boldsymbol{j}$ 和 $\boldsymbol{j} + \boldsymbol{k}$ 都平行且与 x 轴的交点为 $(2,0,0)$ 的平面的笛卡儿方程.
15. 求在以 $3x + y + z = 5$, $3x + y + 5z = 7$, $x - y + 3z = 3$ 为笛卡儿方程的三个平面的交集中的所有点.
16. 证明分别以线性无关的三个向量为法向量的三个平面相交于唯一的一点.

17. 若向量 $\boldsymbol{A}$ 平行于平面 M, 则称任意以 $\boldsymbol{A}$ 为方向向量的直线平行于 M. 若有一直线包含点 $(1,2,3)$ 且与分别以 $x+2y+3z=4$ 和 $2x+3y+4z=5$ 为笛卡儿方程的两个平面都平行, 求该直线的一个向量参数方程.
18. 在 $\mathbb{R}^n$ 中, 给定平面 M 和不平行于 M 的直线 L, 证明它们的交 $L\cap M$ 最多包含一点. 当 $n=3$ 时, 证明 $L\cap M$ 恰好包含一点.
19. (a) 证明点 (x_0,y_0,z_0) 到以 $ax+by+cz+d=0$ 为笛卡儿方程的平面的距离等于

$$\frac{|ax_0+by_0+cz_0+d|}{(a^2+b^2+c^2)^{1/2}}.$$

    (b) 求在以 $5x-14y+2z+9=0$ 为笛卡儿方程的平面上离点 $\boldsymbol{Q}=(-2,15,-7)$ 最近的点 $\boldsymbol{P}$.
20. 求一个平行于以 $2x-y+2z+4=0$ 为笛卡儿方程的平面的平面, 使得点 $(3,2,-1)$ 到这两个平面的距离相等.
21. (a) 若点 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 确定一个平面, 证明点 $\boldsymbol{Q}$ 到该平面的距离等于 $|(\boldsymbol{Q}-\boldsymbol{A})\cdot(\boldsymbol{B}-\boldsymbol{A})\times(\boldsymbol{C}-\boldsymbol{A})|/\|(\boldsymbol{B}-\boldsymbol{A})\times(\boldsymbol{C}-\boldsymbol{A})\|$.
    (b) 计算当 $\boldsymbol{Q}=(1,0,0)$, $\boldsymbol{A}=(0,1,1)$, $\boldsymbol{B}=(1,-1,1)$, $\boldsymbol{C}=(2,3,4)$ 时 $\boldsymbol{Q}$ 到这个平面的距离.
22. 若 $\mathbb{R}^3$ 中两个平面 M 和 M' 不平行, 证明它们的交 $M\cap M'$ 是一条直线.
23. 求平行于 $\boldsymbol{j}$ 且过分别以 $x+2y+3z=4$ 和 $2x+y+z=2$ 为笛卡儿方程的平面的交的平面的笛卡儿方程.
24. 求平行于 $3\boldsymbol{i}-\boldsymbol{j}+2\boldsymbol{k}$ 且包含分别以 $x+y=3$ 和 $2y+3z=4$ 为笛卡儿方程的平面的交的所有点的平面的笛卡儿方程.

2.19 二次曲线

一条直线的运动轨迹构成三维空间中的一个曲面, 若运动直线 G 与一条给定直线 A 交于固定点 P 且与 A 的夹角恒为 θ, 其中 $0<\theta<\frac{1}{2}\pi$, 我们称这样生成的曲面为**直立圆锥** (right circular cone), 称直线 G 为这个圆锥的**母线** (generator), 称 A 为它的**轴** (axis), 称 P 为它的**顶点**(vertex). 图 2.9 中每一个圆锥都有一条竖直的轴. 顶点将圆锥分为上下两部分, 这两部分都称作它的**叶**(nappe). 用不过原点的平面切割圆锥得到的曲线称为**圆锥截线**或**二次曲线**. 如图 2.9 所示, 若割平面与圆锥上一条过顶点的直线平行, 那么我们称得到的二次曲线为**抛物线** (parabola), 否则根据割平面与圆锥的一个叶还是两个叶相交, 我们分别称得到的二次曲线为**椭圆** (ellipse) 或**双曲线** (hyperbola). 双曲线有两个**分支** (branch), 它在两个叶上各有一个分支.

纯数学和应用数学中的许多重要发现都和二次曲线有关. Appolonius 在公元前 3 世纪时关于二次曲线的研究成果是古希腊几何中最伟大的成就之一. 近两千年后, Galileo 发现从塔顶水平发射的抛掷物沿着一条抛物线的路径落到地上 (假定忽略空气阻力并可将地面看作平面). 在公元 1600 年前后, Johannes Kepler(1571—1630) 提出行星都是沿着椭圆轨道运动的, 这是天文学史上的一个转折点. 大约 80 年后, Issac Newton(1642—1727) 由椭圆行星轨道推演出了万有引力的平方反比律,

这项成果是科学史上最伟大的发现之一. 不仅行星和卫星的轨道是二次曲线, 基本原子粒子的轨道也是二次曲线. 在透镜和镜子的设计以及在建筑学中都要用到二次曲线. 此外还有很多例子表明不论怎样强调二次曲线的重要性都不为过.

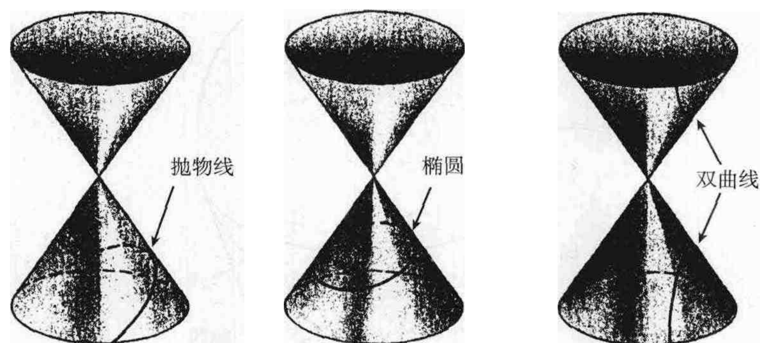


图 2.9 二次曲线

除了刚才我们所说的定义, 二次曲线还有其他等价的定义, 其中之一用到一类称为**焦点** (focus) 的特殊点. 一个椭圆可定义为平面上到两个给定点 F_1 和 F_2(焦点) 的距离 d_1 和 d_2 之和为常数的所有点的集合 (见图 2.10). 若两个焦点重合, 那么这个椭圆化为一个圆.

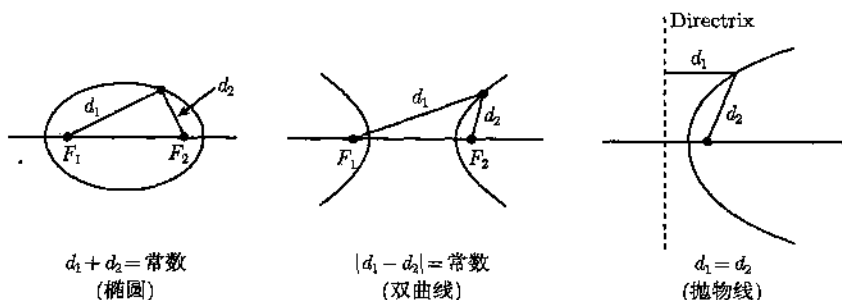


图 2.10 由焦点定义的二次曲线

双曲线是到两个焦点的距离之差 $|d_1 - d_2|$ 为常数的所有点的集合.

抛物线是平面上到一个给定点 F(焦点) 的距离与到一条给定直线的距离相等的所有点的集合. 这条直线称为它的**准线**(directrix).

有一个简单又漂亮的论证表明椭圆的焦点性质是我们将椭圆定义为圆锥的截线的必然结果. 这个结论的证明称为**冰激淋蛋筒证明** (ice-cream cone proof), 它是由比利时数学家 G. F. Dandelin(1794—1847) 在 1822 年发现的. 这个证明使用了如图 2.11 所示的与割平面和圆锥都相切的两个球 S_1 和 S_2, 这两个球与圆锥分别

交于相互平行的圆 C_1 和 C_2，我们将证明两个球与割平面的交点 F_1 和 F_2 是椭圆的两个焦点.

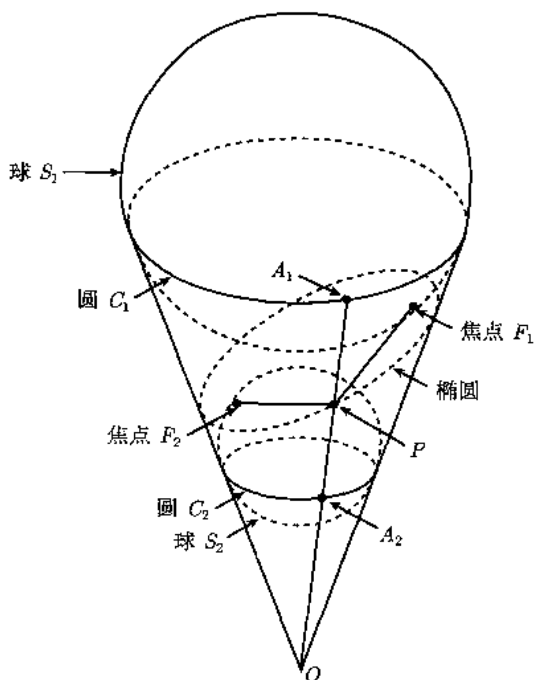


图 2.11　冰激淋蛋筒证明

令 P 为椭圆上任意一点，我们需要证明 $\|\overrightarrow{PF_1}\|+\|\overrightarrow{PF_2}\|$ 是与 P 无关的常数. 为此我们画一条过 O 和 P 的直线并令 A_1 和 A_2 分别为该直线与圆 C_1 和 C_2 的交点，那么 $\overrightarrow{PF_1}$ 和 $\overrightarrow{PA_1}$ 与球 S_1 相切且都过点 P，所以 $\|\overrightarrow{PF_1}\|=\|\overrightarrow{PA_1}\|$. 同理可得 $\|\overrightarrow{PF_2}\|=\|\overrightarrow{PA_2}\|$. 所以我们有

$$\|\overrightarrow{PF_1}\|+\|\overrightarrow{PF_2}\|=\|\overrightarrow{PA_1}\|+\|\overrightarrow{PA_2}\|.$$

又因为 $\|\overrightarrow{PA_1}\|+\|\overrightarrow{PA_2}\|=\|\overrightarrow{A_1A_2}\|$，所以它等于相互平行的圆 C_1 和 C_2 之间沿着圆锥面的距离，这个距离与 P 无关. 所以 F_1 和 F_2 如命题所述是该椭圆的两个焦点.

可将此证明稍作修改来证明抛物线和双曲线的焦点性质. 若考虑双曲线，那么证明会用到分别在圆锥的两个叶中的球；若考虑抛物线，那么证明只会用到一个球它和割平面在焦点 F 处相切. 这个球与圆锥交于一个圆，而这个圆所在平面与割平面的交线就是该抛物线的准线. 有了这些提示，读者可自行证明抛物线和双曲线的焦点性质是将它们定义为圆锥的截线的必然结果.

2.20 二次曲线的离心率

二次曲线另一个标志性性质与离心率概念有关. 二次曲线可定义为一个在平面上运动的点的轨迹, 这个点与一个给定点的距离和它与一条给定直线的距离之比为常数. 这个比值称为该曲线的**离心率** (eccentricity) 并记作 e(不要将它与 Euler 常数 $\mathrm{e} = 2.718\,28\cdots$ 混淆). 若 $0 < e < 1$, 则该曲线是椭圆; 若 $e = 1$, 那么它是抛物线; 若 $e > 1$, 那么它是双曲线. 定义中的给定点称为焦点, 给定直线称为准线.

因为这个定义让我们可同时处理全部三类二次曲线并使向量方法可行, 所以我们将这个定义作为研究二次曲线的基础. 在该定义中, 我们假定所有的点和直线都在同一平面上.

定义 *给定一条直线 L, 一个不在 L 上的点 $\boldsymbol{F}$ 和一个正数 e, 令 $d(\boldsymbol{X}, L)$ 表示点 $\boldsymbol{X}$ 到 L 的距离, 那么满足关系*

$$\|\boldsymbol{X} - \boldsymbol{F}\| = ed(\boldsymbol{X}, L) \tag{2.20}$$

的所有点 $\boldsymbol{X}$ 组成的集合称作离心率为 e 的二次曲线. 若 $e < 1$, 那么称该二次曲线为椭圆; 若 $e = 1$, 那么称它为抛物线; 若 $e > 1$, 那么称它为双曲线.

若 $\boldsymbol{N}$ 是垂直于 L 的向量, $\boldsymbol{P}$ 为 L 上任意点, 则任意点 $\boldsymbol{X}$ 到 L 的距离 $d(\boldsymbol{X}, L)$ 满足 2.5 节中导出的公式:

$$d(X, L) = \frac{|(x - P) \cdot N|}{\|N\|}.$$

当 $\boldsymbol{N}$ 的长度为 1 时, 上式简化为 $d(\boldsymbol{X}, L) = |(\boldsymbol{X} - \boldsymbol{P}) \cdot \boldsymbol{N}|$, 并且二次曲线的基本方程 (2.20) 化为

$$\|\boldsymbol{X} - \boldsymbol{F}\| = e|(\boldsymbol{X} - \boldsymbol{P}) \cdot \boldsymbol{N}|. \tag{2.21}$$

直线 L 将平面分成两部分, 根据 $\boldsymbol{N}$ 的选择, 我们分别称这两部分为"正的"和"负的". 若 $(\boldsymbol{X} - \boldsymbol{P}) \cdot \boldsymbol{N} > 0$, 那么我们称 X 在**正半平面** (positive half-plane) 上; 若 $(\boldsymbol{X} - \boldsymbol{P}) \cdot \boldsymbol{N} < 0$, 那么我们称 X 在**负半平面** (nagetive half-plane) 上. 当 $\boldsymbol{X}$ 在直线 L 上时, 我们有 $(\boldsymbol{X} - \boldsymbol{P}) \cdot \boldsymbol{N} = 0$. 图 2.12 中的法向量 $\boldsymbol{N}$ 决定了直线 L 右边的点在正半平面上, L 左边的点在负半平面上.

如图 2.12 所示, 现在我们将焦点 $\boldsymbol{F}$ 放在负半平面上, 并令 $\boldsymbol{P}$ 为 L 上离 $\boldsymbol{F}$ 最近的点, 那么 $\boldsymbol{P} - \boldsymbol{F} = d\boldsymbol{N}$, 其中 $|d| = \|\boldsymbol{P} - \boldsymbol{F}\|$ 为焦点到准线的距离. 由于 $\boldsymbol{F}$ 在负半平面上, 我们有 $(\boldsymbol{F} - \boldsymbol{P}) \cdot \boldsymbol{N} = -d < 0$, 所以 d 为正. 用 $\boldsymbol{F} + d\boldsymbol{N}$ 代替式 (2.21) 中的 $\boldsymbol{P}$, 那么我们可得下述定理.

定理 2.17 *设 C 是离心率为 e、焦点为 $\boldsymbol{F}$、准线为 L 的二次曲线, 其中 $\boldsymbol{F}$ 到 L 的距离为 d. 若 $\boldsymbol{N}$ 是一个长度为 1 且与 L 垂直的向量, 并且 $\boldsymbol{F}$ 在由 $\boldsymbol{N}$ 决定的负半平面上, 那么 C 由所有满足方程*

$$\|\boldsymbol{X} - \boldsymbol{F}\| = e|(\boldsymbol{X} - \boldsymbol{F}) \cdot \boldsymbol{N} - d|. \tag{2.22}$$

的点 $\boldsymbol{X}$ 组成. □

图 2.12 给出了此定理的几何意义.

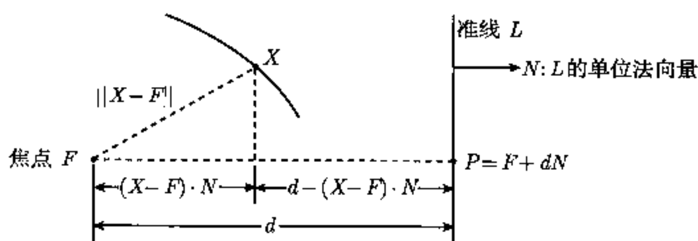


图 2.12 离心率为 e 的二次曲线是满足 $\|\boldsymbol{X}-\boldsymbol{F}\|=e|(\boldsymbol{X}-\boldsymbol{F})\cdot\boldsymbol{N}-d|$ 的所有点 $\boldsymbol{X}$ 的集合

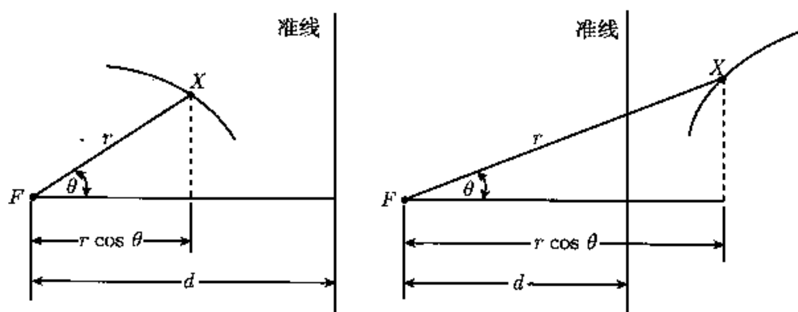
2.21 二次曲线的极坐标方程

当用特殊点作为焦点时, 我们可简化定理 2.17 中的方程. 例如, 若焦点为原点, 那么该方程化为

$$\|\boldsymbol{X}\|=e|\boldsymbol{X}\cdot\boldsymbol{N}-d|. \tag{2.23}$$

这种形式在用极坐标表示 $\boldsymbol{X}$ 时特别有用. 如图 2.13 所示, 假定准线 L 竖直并令 $\boldsymbol{N}=\boldsymbol{i}$. 若 X 的极坐标为 r 和 θ, 那么我们有 $\|\boldsymbol{X}\|=r$, $\boldsymbol{X}\cdot\boldsymbol{N}=r\cos\theta$, 所以方程 (2.23) 化为

$$r=e|r\cos\theta-d|. \tag{2.24}$$



(a) $r\cos\theta<d$ 在椭圆、抛物线和双曲线的左支上　(b) 在双曲线的右支上 $r\cos\theta>d$

图 2.13 极坐标方程为 $r=e|r\cos\theta-d|$ 的二次曲线, 它的焦点 F 为原点, 准线在原点右边

如果 $\boldsymbol{X}$ 在准线的左边, 如图 2.13a 所示, 那么 $r\cos\theta<d$, 于是 $|r\cos\theta-d|=d-r\cos\theta$, 于是方程 (2.24) 化为 $r=e(d-r\cos\theta)$. 对 r 求解即得

$$r=\frac{ed}{e\cos\theta+1}. \tag{2.25}$$

如果 $\boldsymbol{X}$ 在准线的右边, 我们有 $r\cos\theta > d$, 所以方程 (2.24) 化为

$$r = e(r\cos\theta - d),$$

对 r 求解即得

$$r = \frac{ed}{e\cos\theta - 1}. \tag{2.26}$$

由于 $r > 0$, 所以由方程 (2.26) 可知 $e > 1$. 换句话说, 只有双曲线包含准线右边的点, 即我们证明了下述定理, 它的几何意义如图 2.13 所示.

定理 2.18 *设 C 是离心率为 e 的二次曲线且它的焦点 $\boldsymbol{F}$ 为原点, 它的准线 L 为一条在 $\boldsymbol{F}$ 右边的竖直线且与 F 的距离为 d, 如果 $0 < e \leqslant 1$, 那么 C 是椭圆或抛物线, C 上的所有点都在 L 的左边并满足极坐标方程*

$$r = \frac{ed}{e\cos\theta + 1}. \tag{2.27}$$

如果 $e > 1$, 那么 C 是双曲线, 它在 L 的两边各有一个分支, 左支上的点满足式 (2.27), 右支上的点满足

$$r = \frac{ed}{e\cos\theta - 1}. \tag{2.28}$$

□

当准线在其他位置时, 二次曲线的极坐标方程将在下面的一组习题中讨论.

2.22 习 题

1. 若 $\boldsymbol{F}$ 在由 $\boldsymbol{N}$ 决定的正半平面上, 证明定理 2.17 中的方程 (2.22) 必须用方程

$$\|\boldsymbol{X} - \boldsymbol{F}\| = e|(\boldsymbol{X} - \boldsymbol{F}) \cdot \boldsymbol{N} + d|$$

替换.

2. 设 C 为一离心率为 e 的二次曲线, 它的焦点 $\boldsymbol{F}$ 为原点, 它的准线 L 为一条在 $\boldsymbol{F}$ 左边的竖直线且与 $\boldsymbol{F}$ 的距离为 d.

   (a) 若 C 是椭圆或抛物线, 证明 C 上所有点都在 L 的右边且满足极坐标方程

$$r = \frac{ed}{1 - e\cos\theta}.$$

   (b) 若 C 是双曲线, 证明右支上的点都满足 (a) 中给出的方程, 左支上的点都满足 $r = -ed/(1 + e\cos\theta)$. 注意此时 $1 + \cos\theta$ 为负.

3. 若一条二次曲线的焦点为原点, 它的准线是一条在 F 上方的水平直线且与 F 的距离为 d, 证明该曲线上的所有点都满足将 $\sin\theta$ 代替定理 2.18 中的所有方程中的 $\cos\theta$ 后所得的极坐标方程. 若该曲线的准线为在焦点下方的水平线, 那么它的极坐标方程是什么?

习题 4~9 都给出了一条焦点 $\boldsymbol{F}$ 为原点且准线为 $\boldsymbol{F}$ 右边的一条竖直线的二次曲线的极坐标方程, 计算各条曲线的离心率 e 和焦点与准线的距离 d, 并画出每条曲线与它的焦点和准线的关系的简图.

4. $r = \frac{2}{1+\cos\theta}$.

5. $r = \frac{3}{1+\frac{1}{2}\cos\theta}$.

6. $r = \frac{6}{3+\cos\theta}$.

7. $r = \frac{1}{-\frac{1}{2}+\cos\theta}$.

8. $r = \frac{4}{1+2\cos\theta}$.

9. $r = \frac{4}{1+\cos\theta}$.

习题 10~12 都给出了一条焦点 $\boldsymbol{F}$ 为原点的二次曲线的离心率 e 和准线的笛卡儿方程, 计算每条曲线的焦点与准线的距离 d, 并求该二次曲线的极坐标方程. 若二次曲线是双曲线, 给出它每个分支的极坐标方程. 画出各条曲线与它的焦点和准线间关系的简图.

10. $e=\frac{1}{2}$; 准线: $3x+4y=25$.
11. $e=1$; 准线: $4x+3y=25$.
12. $e=2$; 准线: $x+y=1$.
13. 一颗彗星在以太阳为焦点的抛物线轨道上运行, 当它在距太阳 10^8 英里处时, 一个从焦点到该彗星的向量与以焦点为起点且垂直于准线的单位向量 $\boldsymbol{N}$ 的夹角为 $\pi/3$, 而且此焦点在由 $\boldsymbol{N}$ 确定的负半平面中.
    (a) 以焦点为原点, 求此轨道的极坐标方程并计算彗星到太阳的最短距离.
    (b) 当焦点在由 $\boldsymbol{N}$ 确定的正半平面中时求解 (a).

2.23 一般二次曲线的笛卡儿方程

对离心率为 e 的一般二次曲线, 基本方程 (2.22) 可写成以下形式:

$$\|\boldsymbol{X}-\boldsymbol{F}\|=e|(\boldsymbol{X}-\boldsymbol{F})\cdot\boldsymbol{N}-d|=e|\boldsymbol{X}\cdot\boldsymbol{N}-\boldsymbol{F}\cdot\boldsymbol{N}-d|=|e\boldsymbol{X}\cdot\boldsymbol{N}-a|,\tag{2.29}$$

其中 $a=ed+e\boldsymbol{F}\cdot\boldsymbol{N}$. 对上式两边同时取平方可得

$$\|\boldsymbol{X}\|^2-2\boldsymbol{X}\cdot\boldsymbol{F}+\|\boldsymbol{F}\|^2=e^2(\boldsymbol{X}\cdot\boldsymbol{N})^2-2ea\boldsymbol{X}\cdot\boldsymbol{N}+a^2.\tag{2.30}$$

若我们将上式中的向量写成分量形式, 那么我们将得到所有二次曲线都满足的一个笛卡儿方程.

令 $\boldsymbol{X}=(x,y)$, $\boldsymbol{F}=(f_1,f_2)$, $\boldsymbol{N}=(n_1,n_2)$, 将这些分量代入方程 (2.30), 再由 $\|\boldsymbol{N}\|=1$ 可知 $n_1^2+n_2^2=1$, 整理后可得

$$x^2(1-e^2n_1^2)-2e^2n_1n_2xy+y^2(1-e^2n_2^2)+2(aen_1-f_1)x+2(aen_2-f_2)y+G=0,$$

其中 $G=f_1^2+f_2^2-a^2$. 该方程具有形式

$$Ax^2+Bxy+Cy^2+Dx+Ey+G=0,$$

其中

$$A=1-e^2n_1^2,\quad B=-2e^2n_1n_2,\quad C=1-e^2n_2^2.\tag{2.31}$$

换句话说, 任意一条二次曲线总可用一个包含 x, y 的二次项和一次项的笛卡儿方程表示. 接下来我们证明二次项系数 A, B, C 决定二次曲线的**类型** (type), 一次项系数决定曲线在 xy 平面中的位置.

定理 2.19 *每条二次曲线都有一个形如*

$$Ax^2+Bxy+Cy^2+Dx+Ey+G=0\tag{2.32}$$

*的笛卡儿方程. 该曲线的类型由数 $\Delta=4AC-B^2$ 的代数符号确定, 其中 Δ 称为它的***判别式** (discriminant), *Δ 与离心率 e 的关系满足方程*

$$\Delta=4(1-e^2).\tag{2.33}$$

所以当一个二次曲线的判别式为正数、零和负数时, 它分别是椭圆、抛物线和双曲线.

证明 由 $n_1^2+n_2^2=1$ 和公式 (2.31) 中 A, B, C 的表示式, 我们得 $4AC=4(1-e^2)+B^2$, 所以 $4AC-B^2=4(1-e^2)$, 所以若我们引入判别式 $\Delta=4AC-B^2$, 则得到方程 (2.33), 这说明当 $e<1$ 时 $\Delta>0$; 当 $e=1$ 时, $\Delta=0$; 当 $e>1$ 时, $\Delta<0$. □

注 定理 2.19 的逆命题是: 形如方程 (2.32) 的任意笛卡儿方程都表示一条二次曲线. 仅当我们将二次曲线的定义推广到允许退化的二次曲线时, 该命题为真. 例如, 满足 $x^2+y^2+1=0$ 的点集为空集; 满足 $x^2+y^2=0$ 的点集仅包含一点 $(0,0)$, 但是在这两种情况下, 判别式都为正. 如果 A, B, C 全为零, 则判别式也为零, 但是方程 (2.32) 关于 x 和 y 是线性的, 所以它代表一条直线. 若方程 (2.32) 的左边可分解成两个一次因式的乘积, 那么它的图像是两条直线.

2.24 关于原点对称的二次曲线

对一个点集, 如果只要它包含 X, 它就包含 $-X$, 那么我们称该点集关于原点对称. 接下来我们将证明总可以这样来放置给定椭圆或双曲线的一个焦点, 使得该二次曲线关于原点对称. 为此, 我们再次观察方程 (2.30):

$$\|\boldsymbol{X}\|^2-2\boldsymbol{X}\cdot\boldsymbol{F}+\|\boldsymbol{F}\|^2=e^2(\boldsymbol{X}\cdot\boldsymbol{N})^2-2ea\boldsymbol{X}\cdot\boldsymbol{N}+a^2. \tag{2.30}$$

为了使该曲线关于原点对称, 我们需要使当用 $-X$ 代替上式中的 X 时方程仍然成立, 此时我们有

$$\|\boldsymbol{X}\|^2+2\boldsymbol{X}\cdot\boldsymbol{F}+\|\boldsymbol{F}\|^2=e^2(\boldsymbol{X}\cdot\boldsymbol{N})^2+2ea\boldsymbol{X}\cdot\boldsymbol{N}+a^2. \tag{2.34}$$

用式 (2.34) 减式 (2.30) 可得当且仅当 $\boldsymbol{F}\cdot\boldsymbol{X}=ea\boldsymbol{X}\cdot\boldsymbol{N}$, 即当

$$(\boldsymbol{F}-ea\boldsymbol{N})\cdot\boldsymbol{X}=0$$

时, 曲线关于原点对称. 当且仅当因子 $(\boldsymbol{F}-ea\boldsymbol{N})=\boldsymbol{O}$ 时, 曲线上所有点 X 都满足这个方程, 此时

$$\boldsymbol{F}=ea\boldsymbol{N},\quad \text{其中}\quad a=ed+e\boldsymbol{F}\cdot\boldsymbol{N}. \tag{2.35}$$

由 $\boldsymbol{F}=ea\boldsymbol{N}$ 可知 $\boldsymbol{F}\cdot\boldsymbol{N}=ea$, 所以我们有 $a=ed+e^2a$, 即 $a(1-e^2)=ed$. 若 $e=1$, 则这个方程不成立, 这是因为焦点到准线的距离 d 不为零, 这意味着抛物线关于原点不对称. 若 $e\neq 1$, 则我们总可令

$$a=\frac{ed}{1-e^2};\quad \boldsymbol{F}=ea\boldsymbol{N}. \tag{2.36}$$

从而满足条件 (2.35). 注意, 若 $e<1$, 则 $a>0$; 若 $e>1$, 则 $a<0$. 将 $\boldsymbol{F}=ea\boldsymbol{N}$ 代入方程 (2.30) 并整理可得如下定理.

定理 2.20 *令 C 为一条离心率 $e\neq 1$ 的二次曲线而且它的一个焦点 $\boldsymbol{F}$ 与准线 L 的距离为 d. 若 $\boldsymbol{N}$ 是垂直于 L 的单位向量且 $\boldsymbol{F}=ea\boldsymbol{N}$, 其中 $a=ed/(1-e^2)$, 那么 C 是满足方程*

$$\|\boldsymbol{X}\|^2-e^2(\boldsymbol{X}\cdot\boldsymbol{N})^2=a^2(1-e^2) \tag{2.37}$$

的所有点 $\boldsymbol{X}$ 的集合. □

上述方程显示出曲线 C 关于原点的对称性, 这是因为当我们用 $-\boldsymbol{X}$ 代替其中的 $\boldsymbol{X}$ 时, 方程保持不变. 对称性说明椭圆和双曲线都有两个焦点和两条准线, 两个焦点关于曲线中心对称且分别在点 $\pm ae\boldsymbol{N}$ 上, 两条准线也关于曲线的中心对称. 因为椭圆和双曲线都有中心, 所以也称它们为**有心二次曲线** (central conic).

当 $\boldsymbol{X} = \pm a\boldsymbol{N}$ 时, 方程 (2.37) 成立, 我们称这两点为有心二次曲线的顶点. 若曲线为椭圆, 那么连接两个顶点的线段称为曲线的**长轴** (major axis); 若曲线为双曲线, 那么连接两个顶点的线段称为**横截轴** (transverse axis). 易证准线与过两个焦点的直线交于点 $\pm(a/e)\boldsymbol{N}$(见 2.28 节中习题 7).

令 $\boldsymbol{N}'$ 为一个与 $\boldsymbol{N}$ 正交的单位向量, 若 $\boldsymbol{X} = b\boldsymbol{N}'$, 则 $\boldsymbol{X}\cdot\boldsymbol{N} = 0$, 所以当且仅当 $b^2 + e^2a^2 = a^2$ 时, $\boldsymbol{X} = b\boldsymbol{N}'$ 满足方程 (2.37). 此时 $b^2 = a^2(1-e^2)$, 所以 $e < 1$. 连接两点 $X = \pm bN'$ 的线段称为椭圆的**短轴** (minor axis), 其中 $b = a\sqrt{1-e^2}$.

注 在方程 (2.37) 中, 如果我们令 $e = 0$, 那么它化为 $\|\boldsymbol{X}\| = a$, 这是一个圆心在原点、半径为 a 的圆的方程. 从等式 (2.36) 的角度来看, 我们可将这样的圆看作当 $d \to \infty$ 时 $e \to 0$ 并且 $ed \to a$ 的椭圆的极限情况.

2.25 椭圆和双曲线在标准位置时的笛卡儿方程

为了得到椭圆和双曲线的笛卡儿方程, 我们只要把 $\boldsymbol{X}$ 的直角坐标代入向量方程 (2.37) 就行. 当选取 $\boldsymbol{N} = \boldsymbol{i}$ 时, 我们称二次曲线在**标准位置** (standard position) 上, 这意味着两条准线都是竖直的. 现在我们令 $\boldsymbol{X} = (x, y)$, 那么 $\|\boldsymbol{X}\|^2 = x^2 + y^2$ 且 $\boldsymbol{X}\cdot\boldsymbol{N} = x$, 所以方程 (2.37) 化为 $x^2 + y^2 + e^2a^2 = e^2x^2 + a^2$, 即 $x^2(1-e^2) + y^2 = a^2(1-e^2)$, 由此可得方程

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1. \tag{2.38}$$

这个笛卡儿方程既可用来表示椭圆 ($e < 1$) 也可用来表示双曲线 ($e > 1$), 称为它们的笛卡儿方程的**标准型** (standard form). 此时曲线的焦点为 $(ae, 0)$ 和 $(-ae, 0)$, 准线为直线 $x = a/e$ 和 $x = -a/e$. 记住, 当 $e < 1$ 时 $a > 0$, 当 $e > 1$ 时 $a < 0$.

如果 $e < 1$, 我们令 $b = a\sqrt{1-e^2}$ 并将椭圆的笛卡儿方程写成如下标准型

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{2.39}$$

它的两个焦点为 $(c, 0)$ 和 $(-c, 0)$, 其中 $c = ae = \sqrt{a^2+b^2}$. 图 2.14a 给出了椭圆的标准型笛卡儿方程的一个例子.

如果 $e > 1$, 令 $b = |a|\sqrt{e^2-1}$, 并将双曲线的笛卡儿方程写成如下标准型

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \tag{2.40}$$

它的两个焦点为 $(c,0)$ 和 $(-c,0)$, 其中 $c=|a|e=\sqrt{a^2+b^2}$. 图 2.14b 给出了双曲线的标准型笛卡儿方程的一个例子.

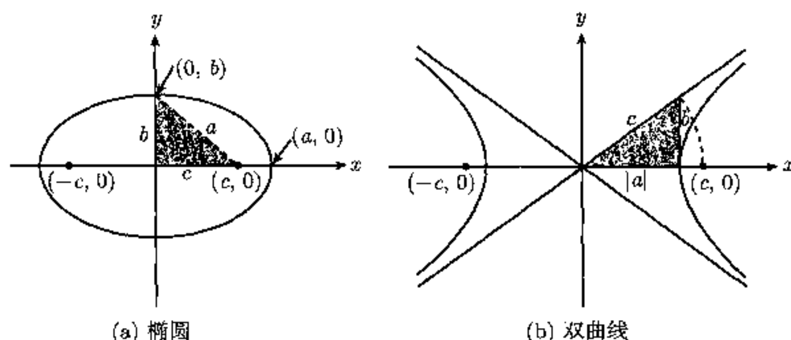


图 2.14 离心率 $e\neq 1$ 的有心二次曲线, 它的焦点为 $(\pm c,0)$, 其中 $c=|a|e$. 图中的三角形表示 a, b, c 的几何关系

注 在方程 (2.40) 中如果用 x 表示 y, 那么我们将得到如下两个关系式:

$$y=\pm\frac{b}{|a|}\sqrt{x^2-a^2}. \tag{2.41}$$

对足够大的正数 x, 数 $\sqrt{x^2-a^2}$ 接近于 x, 所以关系式 (2.41) 的右边接近于 $\pm bx/|a|$. 易证当 $x\to\infty$ 时, $y_1=bx/|a|$ 和 $y_2=b\sqrt{x^2-a^2}/|a|$ 的差趋于 0, 这个差值为

$$y_1-y_2=\frac{b}{|a|}\left(x-\sqrt{x^2-a^2}\right)=\frac{b}{|a|}\frac{x^2-(x^2-a^2)}{x+\sqrt{x^2-a^2}}=\frac{|a|b}{x+\sqrt{x^2-a^2}}<\frac{|a|b}{x}.$$

所以当 $x\to\infty$ 时, $y_1-y_2\to 0$. 我们称直线 $y=bx/|a|$ 为该双曲线的一条**渐近线** (asymptote), 直线 $y=-bx/|a|$ 为它的另一条渐近线. 我们也说该双曲线渐近地趋近于这两条渐近线. 图 2.14b 画出了双曲线的渐近线.

当有心二次曲线的准线不是竖直线时, 它们的笛卡儿方程就会取另一种形式. 例如, 当准线为水平线时, 在方程 (2.37) 中, 我们可设 $\boldsymbol{N}=\boldsymbol{j}$, 因为 $\boldsymbol{X}\cdot\boldsymbol{N}=\boldsymbol{X}\cdot\boldsymbol{j}=y$, 所以我们所得的笛卡儿方程是将方程 (2.38) 中的 x 和 y 互换后所得到的方程. 此时有心二次曲线的笛卡儿方程的标准型为

$$\frac{y^2}{a^2}+\frac{x^2}{a^2(1-e^2)}=1. \tag{2.42}$$

若我们用向量 $\boldsymbol{X}_0=(x_0,y_0)$ 与一条中心在原点的有心二次曲线上的每一点相加, 则得到一条新的有心二次曲线, 这条新曲线的中心为 (x_0,y_0). 此时可用 $x-x_0$ 和 $y-y_0$ 分别代替式 (2.35) 或式 (2.39) 中的 x 和 y 从而获得新曲线的笛卡儿方程.

2.26 抛物线的笛卡儿方程

为了获得抛物线的笛卡儿方程, 我们再来考察当 $e=1$ 时的基本方程 (2.20). 取准线为竖直线 $x=-c$ 并令焦点为 $(c,0)$, 若 $\boldsymbol{X}=(x,y)$, 那么我们有 $\boldsymbol{X}-\boldsymbol{F}=(x-c,y)$, 方程 (2.20) 化为 $(x-c)^2+y^2=|x+c|^2$, 可将此式化简为标准型

$$y^2=4cx. \tag{2.43}$$

我们称焦点到准线的垂直线段的中点 (在图 2.15 中为原点) 为该抛物线的**顶点** (vertex), 称过顶点和焦点的直线为它的**轴** (axis). 抛物线关于它的轴对称. 如图 2.15 所示, 若 $c>0$, 抛物线在 y 轴的右边; 若 $c<0$, 那么抛物线在 y 轴的左边.

若抛物线的焦点在 y 轴上的 $(0,c)$ 处且它的准线为水平线 $y=-c$, 那么它的笛卡儿方程的标准形化为

$$x^2=4cy.$$

如图 2.16 所示, 当 $c>0$ 时, 抛物线开口向上; 当 $c<0$ 时, 抛物线开口向下.

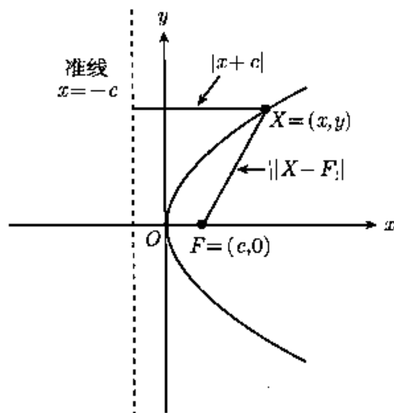


图 2.15 抛物线 $y^2=4cx$

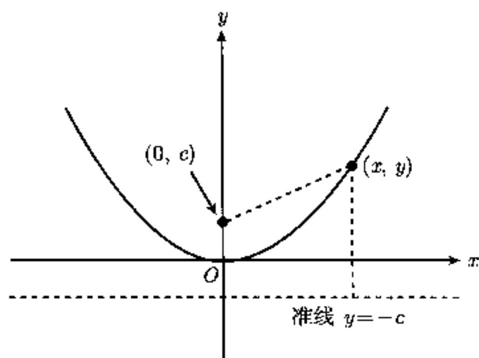


图 2.16 抛物线 $x^2=4cy$

当我们将图 2.15 中的抛物线平移使它的顶点为 (x_0,y_0) 时, 它对应的笛卡儿方程变为

$$(y-y_0)^2=4c(x-x_0).$$

此时抛物线的焦点为 (x_0+c,y_0), 它的准线为直线 $x=x_0-c$, 它的轴为直线 $y=y_0$.

类似地, 当平移图 2.16 中的抛物线使它的顶点为 (x_0,y_0) 时, 它的笛卡儿方程变为

$$(x-x_0)^2=4c(y-y_0),$$

此时抛物线的焦点为 (x_0,y_0+c), 直线 $y=y_0-c$ 是它的准线, 直线 $x=x_0$ 是它的轴.

读者可能会觉得证明抛物线没有渐近线是一件非常有趣的事.

2.27 习　题

习题 1~6 中的每个方程都表示一个椭圆, 求每一个椭圆的中心、焦点和顶点的坐标以及各椭圆的离心率, 并画出各椭圆的简图.

1. $\frac{x^2}{100}+\frac{y^2}{36}=1$.
2. $\frac{y^2}{100}+\frac{x^2}{36}=1$.
3. $\frac{(x-2)^2}{16}+\frac{(y+3)^2}{9}=1$.
4. $9x^2+25y^2=25$.
5. $4y^2+3x^2=1$.
6. $\frac{(x+1)^2}{16}+\frac{(y+2)^2}{25}=1$.

在习题 7~12 中, 求满足所列条件的椭圆的标准型笛卡儿方程并画出各椭圆的简图.

7. 中心为 $(0,0)$, 一个焦点为 $(\frac{3}{4},0)$, 一个顶点为 $(1,0)$.
8. 中心为 $(-3,4)$, 半轴长分别为 4 和 3, 长轴平行于 x 轴.
9. 中心为 $(-3,4)$, 半轴长分别为 4 和 3, 长轴平行于 y 轴.
10. 顶点为 $(-1,2)$, $(-7,2)$, 短轴长为 2.
11. 顶点为 $(3,-2)$, $(13,-2)$, 焦点为 $(4,-2)$, $(12,-2)$.
12. 中心为 $(2,1)$, 长轴平行于 x 轴, 椭圆过点 $(6,1)$ 和 $(2,3)$.

习题 13~18 中的每个方程都表示一条双曲线, 计算每一条双曲线的中心、焦点和顶点的坐标以及各双曲线的离心率, 画出各双曲线的简图并指出它们的渐近线的位置.

13. $\frac{x^2}{100}-\frac{y^2}{64}=1$.
14. $\frac{y^2}{100}-\frac{x^2}{64}=1$.
15. $\frac{(x+3)^2}{4}-(y-3)^2=1$.
16. $9x^2-16y^2=144$.
17. $4x^2-5y^2+20=0$.
18. $\frac{(x-1)^2}{4}-\frac{(y+2)^2}{9}=1$.

在习题 19~23 中, 求满足所列条件的双曲线的标准形笛卡儿方程, 画出各双曲线的简图并指出它们的渐近线的位置.

19. 中心为 $(0,0)$, 一个焦点为 $(4,0)$, 一个顶点为 $(2,0)$.
20. 焦点为 $(0,\pm\sqrt{2})$, 顶点为 $(0,\pm 1)$.
21. 顶点为 $(\pm 2,0)$, 渐近线为 $y=\pm 2x$.
22. 中心为 $(-1,4)$, 一个焦点为 $(-1,2)$, 一个顶点为 $(-1,3)$.
23. 中心为 $(2,-3)$, 横截轴平行于一条坐标轴, 双曲线过点 $(3,-1)$ 和 $(-1,0)$.
24. 某双曲线的渐近线为直线 $2x-y=0$ 和 $2x+y=0$ 并过点 $(3,-5)$, 求该双曲线的笛卡儿方程.

习题 25~30 中的每个方程都表示一条抛物线, 计算它们的顶点坐标, 准线方程, 轴的方程并画出各抛物线的简图

25. $y^2=-8x$.
26. $y^2=3x$.
27. $(y-1)^2=12x-6$.
28. $x^2=6y$.
29. $x^2+8y=0$.
30. $(x+2)^2=4y+9$.

在习题 31~36 中, 求满足所列条件的抛物线的标准型笛卡儿方程并画出各抛物线的简图.

31. 焦点为 $(0,-\frac{1}{4})$, 准线的方程为 $y=\frac{1}{4}$.
32. 顶点为 $(0,0)$, 准线的方程为 $x=-2$.
33. 顶点为 $(-4,3)$, 焦点为 $(-4,1)$.
34. 焦点为 $(3,-1)$, 准线的方程为 $x=\frac{1}{2}$.
35. 轴平行于 y 轴, 抛物线过点 $(0,1)$, $(1,0)$, $(2,0)$.

36. 轴平行于 x 轴, 顶点为 $(1,3)$, 抛物线过点 $(-1,-1)$.
37. 由焦点的定义出发, 直接求焦点为原点、准线为直线 $2x+y=10$ 的抛物线的笛卡儿方程.
38. 求焦点为原点、准线为直线 $x+y+1=0$ 的抛物线的笛卡儿方程.
39. 求由与点 $(0,2)$ 的距离是与直线 $y=8$ 的距离的一半的所有点 (x,y) 组成的二次曲线的笛卡儿方程.
40. 求过原点、渐近线为直线 $y=2x+1$ 和 $y=-2x+3$ 的双曲线的笛卡儿方程.

2.28 关于二次曲线的综合性习题

1. (a) 对每一个 $p>0$, 方程 $px^2+(p+2)y^2=p^2+2p$ 都表示一个椭圆, 求 (用 p 表示) 它的离心率和焦点坐标.

   (b) 求与 (a) 中的椭圆焦点相同且离心率为 $\sqrt{3}$ 的双曲线的笛卡儿方程.
2. 在 2.24 节中, 我们证明了关于原点对称的二次曲线满足方程 $\|\boldsymbol{X}-\boldsymbol{F}\|=|e\boldsymbol{X}\cdot\boldsymbol{N}-a|$, 其中 $a=ed+e\boldsymbol{F}\cdot\boldsymbol{N}$. 用此关系证明: 若二次曲线为椭圆, 那么我们有 $\|\boldsymbol{X}-\boldsymbol{F}\|+\|\boldsymbol{X}+\boldsymbol{F}\|=2a$, 也就是说, 椭圆上任意点到两个焦点的距离之和为常数.
3. 类似于习题 2, 证明在双曲线的每一个分支上, 差 $\|\boldsymbol{X}-\boldsymbol{F}\|-\|\boldsymbol{X}+\boldsymbol{F}\|$ 为常数.
4. (a) 证明相似变换 (用 tx 和 ty 分别替换 x 和 y) 将一个中心在原点的椭圆变为一个与原椭圆离心率相同的另一个椭圆, 也就是说, 相似的椭圆的离心率相等.

   (b) 证明 (a) 的逆命题: 若两个同心 (concentric) 椭圆的离心率相等且长轴在同一条直线上, 那么可通过相似变换将其中的一个变为另一个.

   (c) 对双曲线叙述并证明与 (a) 和 (b) 相应的结论.
5. (a) 证明对一条抛物线作相似变换 (见习题 4) 将得到另一条抛物线.

   (b) 求相似于 $y=x^2$ 的所有抛物线.
6. 如果点 P 和点 Q 与某个圆的圆心共线, 该圆圆心不在 P 和 Q 之间且圆心到 P 和 Q 的距离的乘积等于该圆半径的平方, 那么我们称 P 与 Q 关于该圆对称. 当点 Q 取遍直线 $x+2y=5$ 上的所有点时, 求与 Q 关于圆 $x^2+y^2=4$ 对称的点 P 的轨迹.
7. 对一个如定理 2.20 所述的中心为原点的有心二次曲线, 证明它的两条准线与过它的两个焦点的直线相交于点 $\pm(a/e)N$.
8. 证明抛物线没有渐近线.
9. 设离心率为 e、焦点在 x 轴上的二次曲线过原点且准线为竖直线, 证明该曲线上任意点的直角坐标 (x,y) 都满足一个形如 $y^2=(e^2-1)x^2-cx$ 的方程, 其中 c 为与该曲线有关的常数.
10. 设点 $(2,2)$ 为一条离心率为 $e=\sqrt{2}$ 的双曲线的一个焦点, 该曲线有一条准线为直线 $x+y=2$, 证明该双曲线的笛卡儿方程为 $xy=2$.

    需要微积分知识的习题
11. 求使直线 $3x-2y=C$ 与双曲线 $x^2-3y^2=1$ 相切的常数 C.
12. 一条抛物线弧的底边长为 b, 高为 h, 求由该弧和底边围成的区域的面积.
13. 求将抛物线 $y^2=8x$ 和直线 $x=2$ 围成的区域绕 x 轴旋转所得的旋转体的体积.
14. 笛卡儿方程分别为 $y^2=2(x-1)$ 和 $y^2=4(x-2)$ 的两条抛物线围成一个平面区域 R.

    (a) 用积分法求 R 的面积.

    (b) 求将 R 绕 x 轴旋转所得的旋转体的体积.

    (c) 求将 R 绕 y 轴旋转所得的旋转体的体积.
15. 若 a 和 b 为正数, 证明椭圆 $x^2/a^2+y^2/b^2=1$ 所围区域的面积等于 ab 与一个半径为 1 的圆的面积的乘积.

    注: 不需要任何积分计算, 只需要用积分的基本性质即可证明此命题.

16. 求所有满足下面两个条件的正数 A 和 B: $A > B$, 椭圆 $Ax^2 + By^2 = 3$ 所围区域的面积等于椭圆

$$(A + B)x^2 + (A - B)y^2 = 3.$$

所围区域的面积.

17. 已知直线 $x - y + 4 = 0$ 和抛物线 $y^2 = 16x$ 相切, 求它们的交点.

第3章 线性空间

3.1 引 言

在数学中我们会碰到各种各样可彼此相加并且可和实数相乘的数学对象. 首先, 实数本身就是这样的数学对象, 此外实值函数、复数、无穷级数、n 维空间中的向量以及向量值函数等都是这样的对象. 本章将介绍一个基本数学概念——线性空间. 线性空间不仅包含上述对象, 它还包含许多其他特殊对象.

简单地说, 线性空间是一些可进行特定运算 (称为**加法**和**数乘**) 的任意元素的集合. 在线性空间的定义中, 我们既不规定元素的属性, 也不规定它们之间是如何进行运算的, 而是规定这些运算必须具有的性质并将它们作为线性空间的公理. 接下来我们对这些公理进行详细描述.

3.2 线性空间的公理化定义

令 V 表示已知对象的非空集合, 我们称这些对象为**元素**(element). 当集合 V 满足下述十条公理时, 我们称它为**线性空间**(linear space). 我们将这些公理分为三组.

封闭性公理

公理 1(对加法封闭) 对 V 中任意两个元素 x 和 y, 都有 V 中唯一元素与它们对应, 这个元素称作 x 与 y 的和, 记作 $x+y$.

公理 2(对实数的乘法封闭) 对 V 中任意元素 x 和任意实数 a, 都有 V 中唯一元素与它们对应, 这个元素称作 a 与 x 的积, 记作 ax.

关于加法的公理

公理 3(交换律) 对 V 中任意元素 x 和 y, 我们都有 $x+y=y+x$.

公理 4(结合律) 对 V 中任意元素 x, y, z, 我们都有 $(x+y)+z=x+(y+z)$.

公理 5(零元素的存在性) 在 V 中存在一个元素 (记作 O), 使得对 V 中任意 x, 有

$$x+O=x.$$

公理 6(负元的存在性) 对 V 中任意 x, 元素 $(-1)x$ 有性质

$$x+(-1)x=O.$$

关于数乘的公理

公理 7(**结合律**) 对 V 中任意元素 x 以及任意实数 a 和 b, 我们有

$$a(bx) = (ab)x.$$

公理 8(**对 V 中加法的分配律**) 对 V 中任意元素 x 和 y 以及任意实数 a, 我们有

$$a(x + y) = ax + ay.$$

公理 9(**对实数加法的分配律**) 对 V 中任意元素 x 以及任意实数 a 和 b, 有

$$(a + b)x = ax + bx.$$

公理 10(**单位元的存在性**) 对 V 中任意元素 x, 我们有

$$1x = x.$$

为了强调数乘是实数与 V 中元素相乘, 有时我们也将如上定义的线性空间称为**实线性空间** (real linear space). 参与数乘的数也称为纯量. 如果在公理 2 和公理 7~9 中用复数代替实数, 那么所得的代数结构称为**复线性空间** (complex linear space). 实线性空间以实数为纯量, 复线性空间以复数为纯量. 尽管我们主要讨论实线性空间, 但是所得的定理对复线性空间同样成立. 如果不特别指出一个线性空间是实的或是复的, 那么我们认为它既可以是实的, 也可以是复的.

有些作者将线性空间称为**线性向量空间**或简单地称为**向量空间**. 在本书中, 我们仍然仅将以向量为元素的线性空间称为向量空间.

3.3 线性空间的实例

由于我们没有规定元素的任何性质, 我们称由这十条公理所刻画的代数结构为**抽象** (abstract) 线性空间. 当我们确定集合 V 和如何将它的两个元素相加以及如何用纯量去乘它的元素时, 我们就得到线性空间的一个实例. 读者容易验证下述各例满足关于实线性空间的所有公理.

例 1 令 V 为实数全体的集合 $\mathbb{R}$, 再令 $x + y$ 和 ax 分别为普通实数加法和乘法. □

例 2 令 V 为复数全体的集合 $\mathbb{C}$, 定义 $x + y$ 为复数加法, ax 为复数 x 和实数 a 的乘积. 此时尽管 V 中元素是复数, 但是由于它的纯量是实数, 所以它是一个实线性空间. □

例 3 令 V 为 n 元实数组组成的向量空间 $\mathbb{R}^n$, 它的加法和数乘是用通常方式由分量定义的加法和纯量乘法. □

例 4 令 V 为 $\mathbb{R}^n$ 中与一给定非零向量 $\boldsymbol{N}$ 正交的所有向量的集合. 当 $n = 2$ 时, 该线性空间为一条过原点且以 N 为法向量的直线, 当 $n = 3$ 时, 它是一个过原点且以 N 为法向量的平面. □

下面各例中的空间称为**函数空间**. V 中元素为实值函数, 两个函数 f 和 g 的加法定义为:

$$(f+g)(x)=f(x)+g(x)$$

其中 x 为 f 与 g 的定义域的交集中的任意实数. 函数 f 和实数 a 的乘法定义如下: af 在 f 的定义域内的任意 x 上的函数值为 $af(x)$. 零元素是函数值恒为零的函数. 读者不难验证下述集合都是函数空间.

例 5 定义在一个给定区间上的所有函数的集合. □

例 6 所有多项式的集合. 记住多项式是一个对所有实数都有定义的, 具有如下有限和形式的函数 f:

$$f(x)=c_0+c_1x+c_2x^2+\cdots+c_kx^k=\sum_{i=0}^{k}c_ix^i$$

与 x 的不同幂相乘的数 $c_0,c_1,\cdots,c_k$ 称为**系数**. 若 x^k 是系数不为零的 x 的最高次幂, 则称多项式的次数为 k. 如果除 c_0 以外的所有系数都为零, 那么称这个多项式为一个次数为 0 的常数多项式. $c_0=0$ 的常数多项式称为零多项式, 它没有次数. 所有多项式的集合由零多项式、所有零次多项式以及所有一次多项式、所有二次多项式等所组成. □

例 7 所有次数 $\leqslant n$ 的多项式的集合, 其中 n 为给定非负整数 (尽管零多项式没有次数, 但是每当我们考虑这样的集合时, 我们总认为它包含零多项式). 次数等于 n 的所有多项式的集合不是线性空间, 这是因为它关于加法不满足封闭性公理 (次数等于 n 的两个多项式之和的次数不一定为 n). 例如, 考虑二次多项式 x^2+x 和 $-x^2+x$, 它们的和为一次多项式 $2x$. □

例 8 所有满足条件 $f(1)=0$ 的函数 f 的集合. 数 0 在本例中非常重要, 若我们用一个非零数 c 代替 0, 那么此时封闭性公理不满足. □

例 9~12 要求读者具有微积分的有关知识.

例 9 在给定区间上连续的所有函数的集合. 若给定的区间为 $[a,b]$, 我们将这个空间记为 $C(a,b)$. □

例 10 在给定点上可微的所有函数的集合. □

例 11 在给定区间上可积的所有函数的集合. □

例 12 线性微分方程 $y''+ay'+by=0$ 的所有解的集合, 其中 a,b 为给定常数. 数 0 在这里非常重要, 如果用一个非零数代替等号右边的零, 那么微分方程的解集不满足封闭性公理. □

上述各例和其他许多例子都说明, 线性空间概念已渗透到代数、几何和分析学等领域. 当由线性空间的公理推导出一个定理时, 我们立刻就获得了一个对每个具体的线性空间实例都成立的结果. 有时对某一特殊实例的知识会使我们预见或解释其他实例的结果, 并揭示它们之间原本无法了解的关系.

3.4 公理的简单推论

由线性空间的公理易得下述各定理.

定理 3.1 (**零元素的唯一性**) *在任意线性空间中有且仅有一个零元.*

证明 公理 5 告诉我们, 任意线性空间都至少有一个零元. 假定存在两个零元, 设它们为 O_1 和 O_2, 在公理 5 中令 $x = O_1$, $O = O_2$, 我们有 $O_1 + O_2 = O_1$. 类似地, 我们用 $x = O_2$, $O = O_1$, 则有 $O_2 + O_1 = O_2$. 又因为加法满足交换律, 所以我们有 $O_1 + O_2 = O_2 + O_1$, 因此 $O_1 = O_2$. □

定理 3.2 (**负元的唯一性**) *在任意线性空间中, 任意元素都只有一个负元. 即对任意 x, 有且仅有一个 y 使得 $x + y = O$.*

证明 公理 6 告诉我们, 任意 x 都至少有一个负元 $(-1)x$. 假定 x 有两个负元 y_1 和 y_2, 那么 $x + y_1 = O$ 且 $x + y_2 = O$, 用 y_2 加第一个等式两边并由公理 5, 4, 3, 我们有

$$y_2 + (x + y_1) = y_2 + O = y_2,$$

以及

$$y_2 + (x + y_1) = (y_2 + x) + y_1 = O + y_1 = y_1.$$

因此 $y_1 = y_2$; 所以对 x, 元素 $(-1)x$ 是它唯一的负元. □

注 我们将 x 的负元记作 $-x$, y 与 $(-1)x$ 之和称为 y 与 x 的差, 记作 $y - x$.

下述定理列出了在线性空间中进行的初等代数运算必须满足的一系列性质.

定理 3.3 *在一个给定的线性空间中, 设 x 和 y 表示任意元素, a 和 b 表示任意纯量, 那么我们有如下性质.*

(a) $0x = O$.

(b) $aO = O$.

(c) $(-a)x = -(ax) = a(-x)$.

(d) 若 $ax = O$, 则 $a = 0$ 或 $x = O$.

(e) 若 $ax = ay$ 且 $a \neq 0$, 则 $x = y$.

(f) 若 $ax = bx$ 且 $x \neq O$, 则 $a = b$.

(g) $-(x + y) = (-x) + (-y) = -x - y$.

(h) $x + x = 2x$, $x + x + x = 3x$, 一般地, $\sum_{i=1}^{n} x = nx$.

我们证明性质 (a)~(c), 其余性质的证明留给读者作为练习.

(a) 的证明 令 $z = 0x$, 我们要证明 $z = O$. 将 z 与自己相加并由公理 9, 我们有

$$z + z = 0x + 0x = (0 + 0)x = 0x = z.$$

再将 $-z$ 加到上式两边可得 $z = O$.

(b) 的证明 令 $z = aO$, 将 z 与自己相加再由公理 8 即得.

(c) 的证明 令 $z=(-a)x$, 将 z 与 ax 相加并由公理 9, 我们有

$$z+ax=(-a)x+ax=(-a+a)x=0x=O,$$

所以 z 等于 ax 的负元, 即 $z=-(ax)$. 同理, 如果我们将 $a(-x)$ 与 ax 相加, 则由公理 8 和性质 (b) 可得 $a(-x)=-(ax)$. □

3.5 习题

判断习题 1 ~ 20 中的各个集合关于通常的加法以及与实数的纯量乘法是否为线性空间. 对那些不是线性空间的, 指出它们不满足哪些公理. 习题 9 ~ 20 中的函数为实值函数, 在习题 11 ~ 13 中, 所有函数的定义域都包含 0 和 1.

1. $\mathbb{R}^3$ 中满足条件 $z=0$ 的所有向量 (x,y,z).
2. $\mathbb{R}^3$ 中满足条件 $x=0$ 或 $y=0$ 的所有向量 (x,y,z).
3. $\mathbb{R}^3$ 中满足条件 $y=5x$ 的所有向量 (x,y,z).
4. $\mathbb{R}^3$ 中满足条件 $3x+4y=1$ 且 $z=0$ 的所有向量 (x,y,z).
5. $\mathbb{R}^3$ 中是向量 $(1,2,3)$ 的纯量倍数的所有向量 (x,y,z).
6. $\mathbb{R}^3$ 中分量满足如下三个线性方程的所有向量 (x,y,z):
$$a_{11}x+a_{12}y+a_{13}z=0,\quad a_{21}x+a_{22}y+a_{23}z=0,\quad a_{31}x+a_{32}y+a_{33}z=0.$$
7. $\mathbb{R}^3$ 中所有是 $(1,2,3)$ 和 $(2,3,4)$ 的线性组合的向量 (x,y,z).
8. $\mathbb{R}^n$ 中所有是两个给定向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的线性组合的向量.
9. 所有有理函数 (多项式的商).
10. f 的次数 $\leqslant g$ 的次数 (包含 $f=0$ 的情况) 的所有有理函数 f/g.
11. 满足条件 $f(0)=f(1)$ 的所有函数 f.
12. 满足条件 $2f(0)=f(1)$ 的所有函数 f.
13. 满足条件 $f(1)=1+f(0)$ 的所有函数 f.
14. 所有偶函数: 对任意 x, $f(-x)=f(x)$.
15. 所有奇函数: 对任意 x, $f(-x)=-f(x)$.
16. 对任意 x 都有 $f(x)=f(1-x)$ 的所有函数 f.
17. 所有有界函数: 对任意 x, $|f(x)|\leqslant M$, 其中 M 为依赖于 f 的常数.
18. 所有递增函数: 对任意 $x\leqslant y$, $f(x)\leqslant f(y)$.
19. 所有以 2π 为周期的函数: 对任意 x, $f(x+2\pi)=f(x)$.
20. $\sin x$ 和 $\cos x$ 的所有线性组合.
21. 令 V 为所有正实数的集合 $\mathbb{R}^+$, 定义 V 中两个元素 x 与 y 的"和"为它们在通常意义下的积 xy, 并定义 V 中元素 x 与纯量 c 的"积"为 x^c, 证明 V 是线性空间且 1 是它的零元.
22. (a) 证明可由其他公理推导出公理 10.
    (b) 当用公理 6′ 代替公理 6 时, 证明不可由其他公理推导出公理 10, 其中公理 6′ 为: 对 V 中任意 x, 都有 V 中一个元素 y 使得 $x+y=O$.
23. 令 S 为实数的有序对 (x_1,x_2) 的集合, 当 S 的加法和纯量乘法定义如下时, 分别判断 S 是否为线性空间, 如果不是, 指出它违反了哪条公理.
    (a) $(x_1,x_2)+(y_1,y_2)=(x_1+y_1,x_2+y_2)$, $\quad a(x_1,x_2)=(ax_1,0)$.
    (b) $(x_1,x_2)+(y_1,y_2)=(x_1+y_1,0)$, $\quad a(x_1,x_2)=(ax_1,ax_2)$.
    (c) $(x_1,x_2)+(y_1,y_2)=(x_1,x_2+y_2)$, $\quad a(x_1,x_2)=(ax_1,ax_2)$.
    (d) $(x_1,x_2)+(y_1,y_2)=(|x_1+y_1|,|x_2+y_2|)$, $\quad a(x_1,x_2)=(|ax_1|,|ax_2|)$.

24. 证明定理 3.3 中的性质 (d)~(h).

以下习题需要微积分知识.

25. 将加法和纯量乘法分别定义为通常意义下的加法和纯量乘法时，判断下述各实值函数的集合是否为线性空间.
    (a) $[0,1]$ 上满足条件 $\int_0^1 f(x)\mathrm{d}x=0$ 的所有可积函数 f.
    (b) $[0,1]$ 上满足条件 $\int_0^1 f(x)\mathrm{d}x\geqslant 0$ 的所有可积函数 f.
    (c) 满足当 $x\to\infty$ 时, $f(x)\to 0$ 的所有函数 f.
    (d) 满足二阶微分方程 $f''+P(x)f'+Q(x)f=0$ 的所有函数 f, 其中 P 和 Q 为在任意点上都连续的给定函数.

3.6 线性空间的子空间

给定线性空间 V, 令 S 为 V 的一个非空子集, 若 S 关于同样的加法和纯量乘法也是一个线性空间, 那么我们称 S 是 V 的一个**子空间**. 为了判断一个集合是不是线性空间, 我们必须验证它是否满足全部十条公理. 然而下述定理表明, 要判断某线性空间的子集是否也是线性空间, 我们只需要验证它是否满足封闭性公理.

定理 3.4 令 S 为线性空间 V 的非空子集, 那么当且仅当 S 满足封闭性公理时它是 V 的子空间.

证明 若 S 是一个子空间, 那么它满足线性空间的所有公理, 由此可知它也满足封闭性公理.

接下来我们证明当 S 满足封闭性公理时, 它也满足其他各公理. 因为对 V 中所有元素, 加法的交换律 (公理 3) 和结合律 (公理 4) 以及关于纯量乘法的所有公理 (公理 7~10) 都成立, 所以它们对 S 中所有元素也成立. 因此我们只需要证明 S 也满足公理 5 和公理 6, 即 S 中存在零元素以及 S 中每个元素都有负元.

令 x 为 S 中任意元素 (由于 S 非空, 所以它至少包含一个元素), 由公理 2, 对任意纯量 a, ax 都属于 S, 令 $a=0$, 即得 $0x$ 也属于 S. 由定理 3.3(a) 有 $0x=O$, 所以 $O\in S$, 故公理 5 成立. 再令 $a=-1$, 即得 $(-1)x$ 属于 S. 又因为 x 和 $(-1)x$ 都属于 V, 所以 $x+(-1)x=O$, 由此公理 6 成立, 所以 S 是 V 的子空间. □

在下述各例中, V 是所有定义在实轴上的实函数组成的线性空间, S 是 V 的一个子集.

例 1 若 S 是所有偶函数组成的集合, 即对任意 x 都有 $f(-x)=f(x)$ 的函数, 那么 S 满足封闭性公理, 所以 S 是 V 的子空间. □

例 2 若 S 是所有满足 $f(0)=0$ 的函数 f 组成的集合, 那么 S 满足封闭性公理, 所以 S 是 V 的子空间. □

例 3 若 S 是所有满足 $f(0)=1$ 的函数 f 组成的集合, 那么 S 不满足封闭性公理, 所以它不是 V 的子空间. □

例 4 若 S 是所有形如 $f(x)=ae^x+be^{-x}$ 的函数 f 组成的集合, 其中 a 和 b 是常数, 那么 S 满足封闭性公理, 所以 S 是 V 的子空间. 我们称形如 ae^x+be^{-x} 的函数为 e^x 和 e^{-x} 的线性组合. □

定义 令 S 为线性空间 V 的一个非空子集, V 中形如

$$x=\sum_{i=1}^{k}c_ix_i,$$

的元素 x 称为 S 中元素的一个**有限线性组合**, 其中 $x_1,\cdots,x_k$ 属于 S, $c_1,\cdots,c_k$ 为纯量. S 中元素的所有有限线性组合的集合满足封闭性公理, 所以它是 V 的一个子空间, 我们说这个子空间**由 S 生成**, 也称它是由 S **线性生成** (linear span) **的子空间**, 记作 $L(S)$. 若 S 为空集, 我们定义 $L(S)$ 为 $\{O\}$, 即仅由零作为元素的集合.

不同的集合可能会生成相同的子空间. 例如, 下述各向量组都生成空间 $\mathbb{R}^2$: $\{\boldsymbol{i},\boldsymbol{j}\}$, $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{i}+\boldsymbol{j}\}$, $\{O,\boldsymbol{i},-\boldsymbol{i},\boldsymbol{j},-\boldsymbol{j},\boldsymbol{i}+\boldsymbol{j}\}$.

所有次数 $\leqslant n$ 的多项式 $p(t)$ 组成的空间可由 $n+1$ 个多项式组成的集合

$$\{1,t,t^2,\cdots,t^n\},$$

生成, 也可由集合 $\{1,t/2,t^2/3,\cdots,t^n/(n+1)\}$ 或集合 $\{1,(1+t),(1+t)^2,\cdots,(1+t)^n\}$ 生成.

全体多项式组成的空间不可能由只有有限个多项式组成的集合生成, 然而这个空间可由无限多个多项式的集合 $\{1,t,t^2,\cdots\}$ 生成.

由此很自然地引出一系列问题. 例如, 什么样的空间可由有限集生成? 如果一个空间可由有限集生成, 那么生成它的集合至少包含多少个元素? 为了讨论这两个问题以及其他相关问题, 我们引入**线性相关**、**线性无关**、**基**和**维数**等概念. 在第 1 章研究向量空间 $\mathbb{R}^n$ 时, 我们已经接触过这些概念了, 现在我们将它们推广到一般线性空间.

3.7 线性空间的线性相关组和线性无关组

定义 设 S 是由 V 中的若干元素组成的集合, 若存在 S 中有限个不同的元素 $x_1,\cdots,x_k$ 和不全为零的一组纯量 $c_1,\cdots,c_k$, 使得

$$\sum_{i=1}^{k}c_ix_i=O. \tag{3.1}$$

那么我们称 S **线性相关**. 形如 (3.1) 且 c_i 不全为零的等式称为 O 的一个**非平凡表示**.

若集合 S 不是线性相关的, 那么我们说它**线性无关**. 此时, 对任意 S 中不同的元素 $x_1,\cdots,x_k$ 和纯量 $c_1,\cdots,c_k$, 由

$$\sum_{i=1}^{k} c_i x_i = O,$$

必有 $c_1 = c_2 = \cdots = c_k = 0$.

尽管线性相关和线性无关是元素组的性质, 我们也将它们用于元素自身. 例如, 我们将线性无关组的元素称为线性无关元素.

如果 S 是有限集, 则上述定义与第 1 章中给出的 $\mathbb{R}^n$ 空间中的线性相关和线性无关的定义一致. 然而上述定义并不限于有限集. 如果 S 是无限集, 则当且仅当它的所有有限子集都线性无关时, S 线性无关.

例 1 若 $T \subseteq S$ 且 T 线性相关, 那么 S 也线性相关. 这和命题 "线性无关集的所有子集都线性无关" 逻辑等价. □

例 2 若 S 中某个元素是另一个元素的纯量倍数, 那么 S 线性相关. □

例 3 任意包含零元素的集合都线性相关. □

例 4 空集线性无关. □

在第 1 章中我们给出了很多由 n 维空间中的向量组成的线性相关组和线性无关组. 下面给出函数空间中线性相关组和线性无关组的例子. 其中线性空间 V 由定义在实轴上的所有实值函数组成.

例 5 对所有实数 t, 令 $u_1(t) = \cos^2 t$, $u_2(t) = \sin^2 t$, $u_3(t) = 1$, 那么由 Pythagoras 恒等式 $\cos^2 t + \sin^2 t = 1$ 可知 $u_1 + u_2 - u_3 = O$, 所以函数 u_1, u_2, u_3 线性相关.

例 6 对实数 t 和 $k = 0, 1, 2, \cdots$, 令 $u_k(t) = t^k$, 则无限集 $S = \{u_0, u_1, u_2, \cdots\}$ 线性无关. 为了证明这个结论, 只需证明对任意 $n, n+1$ 个多项式 $u_0, u_1, \cdots, u_n$ 都线性无关. 如果这些多项式生成零多项式, 那么对任意实数 t 都有

$$\sum_{k=0}^{n} c_k t^k = 0 \tag{3.2}$$

在式 (3.2) 中令 $t = 0$ 可得 $c_0 = 0$, 再用 t 除等式 (3.2) 的两边并令 $t = 0$ 可知 $c_1 = 0$. 重复这个过程我们知道所有 c_k 都为零, 所以 $u_0, u_1, \cdots, u_n$ 线性无关. □

例 7 (在线性微分方程的研究中需要用到本例.) 设 $a_0, a_1, \cdots, a_n$ 为互不相同的实数, 那么 n 个指数函数

$$u_1(x) = \mathrm{e}^{a_1 x}, \cdots, u_n(x) = \mathrm{e}^{a_n x}$$

线性无关. 我们对 n 用归纳法证明这个结论. 当 $n = 1$ 时, 结论显然成立. 设对 $n-1$ 个指数函数结论成立, 考虑纯量 $c_1, \cdots, c_n$ 使得

$$\sum_{k=1}^{n} c_k \mathrm{e}^{a_k x} = 0 \tag{3.3}$$

令 a_M 为 n 个数 $a_1, \cdots, a_n$ 中最大的一个, 用 $\mathrm{e}^{-a_M x}$ 乘等式 (3.3) 两边可得

$$\sum_{k=1}^{n} c_k e^{(a_k-a_M)x}=0 \tag{3.4}$$

对任意 $k\neq M$, 指数因子 a_k-a_M 都是负数, 所以当我们令 $x\to+\infty$ 时, 等式 (3.4) 中所有 $k\neq M$ 的项都趋于零, 由此可知 $c_M=0$. 所以可将等式 (3.3) 中第 M 项划去, 再由归纳假设可知, 剩下的所有项的系数 c_k 都为零. □

定理 3.5 设 S 是线性空间 V 中由 k 个元素组成的线性无关组 $\{x_1,\cdots,x_k\}$, 令 $L(S)$ 表示由 S 生成的子空间, 那么 $L(S)$ 中任意 $k+1$ 个元素构成的集合都线性相关.

证明 对 S 中元素的个数用归纳法. 首先设 $k=1$, 所以 $S=\{x_1\}$. 因为 S 无关, 所以 $x_1\neq O$. 取 $L(S)$ 中任意两个不同的元素 y_1 和 y_2, 那么它们都是 x_1 的纯量倍数. 设 $y_1=c_1x_1$, $y_2=c_2x_1$, 则 c_1 和 c_2 不全为零, 用 c_2 乘 y_1, c_1 乘 y_2, 并将它们相减可得 $c_2y_1-c_1y_2=O$, 这是 O 的一个非平凡表示, 所以 y_1 和 y_2 线性相关, 因此当 $k=1$ 时定理得证.

现在假定对于 $k-1$ 定理成立, 我们来证明对于 k 它也成立. 取 $L(S)$ 中任意包含 $k+1$ 个元素的集合 $T=\{y_1,y_2,\cdots,y_{k+1}\}$, 我们需要证明 T 线性相关. 对任意 $i=1,2,\cdots,k+1$, 由于 y_i 都属于 $L(S)$, 因此它可写成各个 x_j 的线性组合, 设为

$$y_i=\sum_{j=1}^{k} a_{ij}x_j \tag{3.5}$$

观察与 x_1 相乘的纯量 a_{i1}, 根据它们是否全为零, 分两种情形讨论.

情形 1 对任意 $i=1,2,\cdots,k+1$, 都有 $a_{i1}=0$. 此时表达式 (3.5) 中的和不包含关于 x_1 的项, 所以每个 T 中的 y_i 都在集合 $S'=\{x_2,\cdots,x_k\}$ 的线性生成集中. 又因为 S' 无关且包含 $k-1$ 个元素, 由归纳假设, 定理对 $k-1$ 成立, 于是集合 T 线性相关, 于是对情形 1 定理成立.

情形 2 各纯量 a_{i1} 不全为零. 不妨设 $a_{11}\neq 0$(若有必要, 我们可对 y_i 重新编号使得 $a_{11}\neq 0$). 在表达式 (3.5) 中令 $i=1$ 并用纯量 $c_i=a_{i1}/a_{11}$ 乘等式两边, 那么我们有

$$c_iy_1=a_{i1}x_1+\sum_{j=2}^{k} c_ia_{1j}x_j.$$

将上式与式 (3.5) 相减可得, 对 $i=2,\cdots,k+1$, 有

$$c_iy_1-y_i=\sum_{j=2}^{k}(c_ia_{1j}-a_{ij})x_j.$$

这个等式将 k 个元素 $c_iy_1-y_i$ 都表示为 $k-1$ 个元素 $x_2,\cdots,x_k$ 的线性组合, 由归纳假设, 这 k 个元素 $c_iy_1-y_i$ 必线性相关, 所以有不全为零的纯量 $t_2,\cdots,t_{k+1}$, 使

得

$$\sum_{i=2}^{k+1} t_i(c_i y_1 - y_i) = O,$$

由此可知

$$\left(\sum_{i=2}^{k+1} t_i c_i\right) y_1 - \sum_{i=2}^{k+1} t_i y_i = O.$$

上式将零元素表示为 $y_1, y_2, \cdots, y_{k+1}$ 的一个非平凡线性组合, 所以 $y_1, y_2, \cdots, y_{k+1}$ 必线性相关, 定理得证. □

3.8 基与维数

定义 设 S 为线性空间 V 中有限个元素组成的集合, 若 S 线性无关且生成 V, 则称 S 为 V 的**一组有限基**. 若线性空间 V 仅包含 O 或有一组有限基, 则称它为**有限维空间**, 否则称它为**无限维空间**.

定理 3.6 设 V 是一个有限维线性空间, 那么 V 的所有有限基包含的元素个数都相同.

证明 令 S 和 T 为 V 的两组有限基, 设 S 包含 k 个元素, T 包含 m 个元素. 由于 S 线性无关且生成 V, 由定理 3.5 可知, 任意一个包含 V 中 $k+1$ 个元素的集合都线性相关, 所以任意一个包含 V 中多于 k 个元素的集合都线性相关. 又因为 T 为包含 V 中 m 个元素的线性无关组, 所以 $m \leqslant k$. 同理在上述论证中交换 S 和 T 可知 $k \leqslant m$, 因此 $k = m$. □

定义 若线性空间 V 有一组包含 n 个元素的基, 那么我们称整数 n 为 V 的**维数**, 并将它记作 $n = \dim V$. 若 $V = \{O\}$, 那么我们称它的维数为 0.

例 1 空间 $\mathbb{R}^n$ 的维数为 n. n 个单位坐标向量组成的集合是它的一组基. □

例 2 次数 $\leqslant n$ 的所有多项式构成的线性空间的维数为 $n+1$. 包含 $n+1$ 个多项式的集合 $\{1, t, t^2, \cdots, t^n\}$ 是它的一组基, 任意一个次数 $\leqslant n$ 的多项式都是这些多项式的一个线性组合. □

例 3 所有多项式构成一个无限维线性空间. 尽管多项式的无穷集 $\{1, t, t^2, \cdots\}$ 无关且生成该空间, 但是没有多项式的有限集生成该空间, 所以它没有有限基. □

例 4 (本例需要微分方程的知识.) 线性微分方程 $y'' - 2y' - 3y = 0$ 有两个线性无关的解 $u_1(x) = \mathrm{e}^{-x}$ 和 $u_2(x) = \mathrm{e}^{3x}$ 且它的所有解都是这两个解的线性组合, 因此该方程的解空间有一组包含两个元素的基 $\{u_1, u_2\}$, 于是它的维数为 2. □

定理 3.7 设 V 是一个有限维线性空间且 $\dim V = n$, 那么我们有:

(a) V 中任意一个线性无关组都是 V 的某组基的子集;

(b) 由 V 的任意 n 个线性无关元素组成的集合都是 V 的基.

证明　为了证明 (a), 令 $S=\{x_1,\cdots,x_k\}$ 为 V 中任意一个线性无关组. 如果 $L(S)=V$, 那么 S 是 V 的基. 否则, 存在 V 中不属于 $L(S)$ 的元素 y, 考虑 $\{y\}$ 和 S 的并集 $S'=\{x_1,\cdots,x_k,y\}$, 我们将证明 S' 线性无关. 不然的话, 存在不全为零的纯量 $c_1,\cdots,c_{k+1}$, 使得

$$\sum_{i=1}^{k}c_ix_i+c_{k+1}y=O.$$

又因为 $x_1,\cdots,x_k$ 线性无关, 所以 $c_{k+1}\neq 0$, 因此我们可由上式解出 y, 由此可知 y 属于由 S 生成的子空间, 这与 y 不属于 $L(S)$ 矛盾, 所以 S' 线性无关且包含 $k+1$ 个元素. 如果 $L(S')=V$, 那么 S' 是 V 的基, 且因为 S 是 S' 的子集, (a) 得证. 如果 S' 不是 V 的基, 我们可对 S' 重复上述过程, 从而得到包含 $k+2$ 个元素的新的线性无关组 S'', 如果 S'' 是 V 的基, 那么 (a) 得证, 否则, 我们再重复上述过程, 最终将在有限步内获得 V 的一组基. 不然的话, 我们会得到一个包含 $n+1$ 个元素的线性无关组, 这与定理 3.5 矛盾, 所以 (a) 得证.

为了证明 (b), 令 S 为任意包含 n 个元素的无关集. 由 (a) 可知, S 是 V 的某组基 B 的子集. 又由定理 3.6 可知 B 恰好包含 n 个元素, 所以 $S=B$. □

注　(b) 表明 n 维线性空间中任意 n 个线性无关元素组成的集合都生成该空间.

3.9　分　　量

设 V 为一个 n 维线性空间, 考虑 V 的一组基, 它的元素 $e_1,\cdots,e_n$ 按给定顺序排列, 我们将这样的有序基记为一个 **n 序组** $(e_1,\cdots,e_n)$. 此时 V 中任意元素 x 都可表示为这组基中元素的线性组合

$$x=\sum_{i=1}^{n}c_ie_i. \tag{3.6}$$

等式中的系数构成一个 n 序组 $(c_1,\cdots,c_n)$, 它由 x 唯一确定. 事实上, 如果我们可将 x 表示为同一组基的元素的另一个线性组合, 设这个线性组合为 $x=\sum\limits_{i=1}^{n}d_ie_i$, 用表达式 (3.6) 减此式可得 $\sum\limits_{i=1}^{n}(c_i-d_i)e_i=O$, 又因为同一组基中的元素线性无关, 所以对任意 i, 都有 $c_i-d_i=0$, 因此 $(c_1,\cdots,c_n)=(d_1,\cdots,d_n)$.

由表达式 (3.6) 确定的 n 序组 $(c_1,\cdots,c_n)$ 的各分量称为 x 关于有序基 $(e_1,\cdots,e_n)$ 的分量.

例 1　设 P_n 为次数不大于给定整数 n 的所有实多项式组成的线性空间, 那么该空间的维数为 $n+1$, 集合 $\{1,t,t^2,\cdots,t^n\}$ 是它的基. 如果多项式 $f(t)$ 可表示为 $f(t)=c_0+c_1t+\cdots+c_nt^n$, 那么系数 $c_0,c_1,\cdots,c_n$ 是 f 关于有序基 $(1,t,t^2,\cdots,t^n)$ 的分量. □

例 2　函数 e^x 和 e^{-x} 的所有线性组合组成的集合是一个二维线性空间, $\{e^x, e^{-x}\}$ 是它的基. 双曲函数 $\cosh x = (e^x + e^{-x})/2$ 和 $\sinh x = (e^x - e^{-x})/2$ 是该空间中的元素, 它们关于有序基 (e^x, e^{-x}) 的分量分别为 $(\frac{1}{2}, \frac{1}{2})$ 和 $(\frac{1}{2}, -\frac{1}{2})$. □

3.10 习　题

在习题 1 ~ 10 中, S 表示 $\mathbb{R}^3$ 中分量满足给定条件的所有向量 (x,y,z) 组成的集合, 判断 S 是否为 $\mathbb{R}^3$ 的子空间, 如果是, 求 $\dim S$.

1. $x = 0$.
2. $x + y = 0$.
3. $x + y + z = 0$.
4. $x = y$.
5. $x = y = z$.
6. $x = y$ 或 $x = z$.
7. $x^2 - y^2 = 0$.
8. $x + y = 1$.
9. $y = 2x$ 且 $z = 3x$.
10. $x + y + z = 0$ 且 $x - y - z = 0$.

令 P_n 表示次数不大于给定整数 n 的所有实多项式组成的线性空间, 在习题 11~22 中, 令 S 为 P_n 中满足给定条件的所有多项式 f 组成的集合, 判断 S 是否为 P_n 的子空间, 如果是, 求 $\dim S$. (习题 19~22 需要用到微积分知识.)

11. $f(0) = 0$.
12. $f(0) = f(2)$.
13. $f(0) + f(1) = 0$.
14. f 为偶多项式.
15. f 为奇多项式.
16. 对所有 x, $f(x) \geqslant 0$.
17. f 的二次数 $\leqslant k$, 其中 $k < n$, 或 $f = 0$.
18. f 的次数为 k, 其中 $k < n$, 或 $f = 0$.
19. $f'(0) = 0$.
20. $f'(0) = 0$ 和 $f(0) = 0$.
21. $f''(0) = 0$.
22. $f'(0) + f(0) = 0$.
23. 在所有实多项式 $p(t)$ 构成的线性空间中, 分别描述下列各多项式的子集生成的子空间, 并求它们的维数.
    (a) $\{1, t^2, t^4\}$　(b) $\{1, t^3, t^5\}$　(c) $\{t, t^2\}$　(d) $\{1 + t, (1 + t)^2\}$.
24. 设 V 是由所有定义在实轴上的实值函数组成的线性空间, 判断下列各题中 V 的子集是否线性相关并求它们生成的子空间的维数.
    (a) $\{1, e^{ax}, e^{bx}\}, a \neq b$.
    (b) $\{e^{ax}, xe^{ax}\}$.
    (c) $\{1, e^{ax}, xe^{ax}\}$.
    (d) $\{e^{ax}, xe^{ax}, x^2e^{ax}\}$.
    (e) $\{e^x, e^{-x}, \cosh x\}$.
    (f) $\{\cos x, \sin x\}$.
    (g) $\{\cos^2 x, \sin^2 x\}$.
    (h) $\{1, \cos 2x, \sin^2 x\}$.
    (i) $\{\sin x, \sin 2x\}$.
    (j) $\{e^x \cos x, e^{-x} \sin x\}$.
25. 在本题中, $L(S)$ 表示线性空间 V 的子集 S 生成的子空间, 证明 (a)~(f) 各小题中的结论.
    (a) $S \subseteq L(S)$.
    (b) 如果 $S \subseteq T \subseteq V$ 且 T 是 V 的子空间, 那么 $L(S) \subseteq T$. 这一性质可以叙述为"$L(S)$ 是 V 的包含 S 的最小子空间".
    (c) 设 S 为 V 的子集, 则当且仅当 $L(S) = S$ 时 S 是 V 的子空间.
    (d) 如果 $S \subseteq T \subseteq V$, 那么 $L(S) \subseteq L(T)$.
    (e) 如果 S 和 T 都是 V 的子空间, 那么 $S \cap T$ 也是 V 的子空间.
    (f) 如果 S 和 T 都是 V 的子集, 那么 $L(S \cap T) \subseteq L(S) \cap L(T)$.
    (g) 给出使得 $L(S \cap T) \neq L(S) \cap L(T)$ 的 V 的两个子集 S 和 T.

26. 设 V 为一个有限维线性空间, S 是 V 的子空间, 证明下列各命题.

(a) S 是有限维空间且 $\dim S \leqslant \dim V$.

(b) 当且仅当 $S=V$ 时 $\dim S=\dim V$.

(c) S 的任意一组基都是 V 的某组基的子集.

(d) V 的基未必包含 S 的基.

3.11 内积・欧氏空间・范数

在通常意义上的欧氏几何中, 几何图形的依赖于线段长度和直线间夹角的性质称为**度量性质** (metric property). 在研究 $\mathbb{R}^n$ 时, 我们用点积定义了长度和夹角, 现在我们希望将这些概念推广到更一般的线性空间中去. 为此我们首先将点积推广为线性空间中的**内积** (inner product), 然后我们再用内积来定义长度和夹角.

在第 1 章中, 我们将 $\mathbb{R}^n$ 中两个向量 $\boldsymbol{x}=(x_1,\cdots,x_n)$ 和 $\boldsymbol{y}=(y_1,\cdots,y_n)$ 的点积 $\boldsymbol{x}\cdot\boldsymbol{y}$ 定义为

$$\boldsymbol{x}\cdot\boldsymbol{y}=\sum_{i=1}^{n}x_iy_i \tag{3.7}$$

在一般的线性空间中, 我们将内积记作 (x,y) 而不是 $x\cdot y$, 而且我们不是用特定的公式而是用公理化方法来定义内积, 即我们列出内积需要满足的一组性质并将这些性质看作公理.

定义 设 V 是一个实线性空间, 如果对 V 中任意一对元素 x 和 y, 都存在一个实数 (x,y) 与之对应, 使得对 V 中任意元素 x,y,z 和任意实数 c, 上述对应关系都满足下列公理:

(1) $(x,y)=(y,x)$ (交换性, 对称性),

(2) $(x,y+z)=(x,y)+(x,z)$ (分配性, 线性性),

(3) $c(x,y)=(cx,y)$ (结合性, 齐次性),

(4) 若 $x\neq O$, 则 $(x,x)>0$ (正定性),

则称在 V 上定义了**内积**, 并称定义了内积的实线性空间为**实欧氏空间** (real Euclidean space).

注 在 (3) 中令 $c=0$ 可知, 对任意 y 都有 $(O,y)=0$.

在复线性空间中, 内积 (x,y) 是一个复数且它也满足关于实内积的除对称性外的所有公理, 我们将对称性公理用下述公理代替:

(1′) $(x,y)=\overline{(x,y)}$ (Hermit[①]对称性),

其中 $\overline{(y,x)}$ 表示 (y,x) 的复共轭. 在齐次性公理 (3) 中, 纯量 c 可为任意复数. 由 (3) 和 (1′) 可得关系

(3′) $(x,cy)=\overline{(cy,x)}=\overline{c}\overline{(y,x)}=\overline{c}(x,y)$.

① 为了纪念对代数和分析作出众多贡献的法国数学家 Charles Hermite (1822—1901).

也就是说, 当从内积的第二个因子中提取纯量时, 必须对它取共轭.

我们称定义了内积的复线性空间为**复欧氏空间**(complex Euclidean space), 有时也称它为**酉空间**(unitary space). 1.16 节中简略讨论的复空间 $\mathbb{C}^n$ 就是这样的空间.

尽管我们主要关注实欧氏空间, 但是本章中的定理对复欧氏空间也成立. 当我们只说欧氏空间而不作其他说明时, 它既可以是实欧氏空间也可以是复欧氏空间.

读者可证明下述各例都满足实值内积的所有公理.

例 1 在 $\mathbb{R}^n$ 中令 $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$, 即如等式 (3.7) 定义的通常意义下 x 和 y 的点积. □

例 2 在 $\mathbb{R}^2$ 中定义两个向量 $\boldsymbol{x} = (x_1, x_2)$ 和 $\boldsymbol{y} = (y_1, y_2)$ 的内积为

$$(\boldsymbol{x}, \boldsymbol{y}) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2.$$

本例表明一个给定线性空间中可能有多个内积. □

下面的三个例子要求读者具有积分知识.

例 3 设 $C(a, b)$ 是在区间 $[a, b]$ 上连续的所有实值函数组成的线性空间, 定义两个连续函数 f 和 g 的内积为

$$(f, g) = \int_a^b f(t)g(t)\mathrm{d}t$$

上式和定义 $\mathbb{R}^n$ 中两向量的内积的等式 (3.7) 相似, 函数值 $f(t)$, $g(t)$ 对应于等式 (3.7) 中的分量 x_i, y_i, 积分运算对应于等式 (3.7) 中的求和运算. □

例 4 在例 3 定义的空间 $C(a, b)$ 中, 定义

$$(f, g) = \int_a^b w(t)f(t)g(t)\mathrm{d}t,$$

其中 w 为 $C(a, b)$ 中一个给定的正函数, 函数 w 称为**权函数** (weight function). 在例 3 中, 对任意 t, 都有 $w(t) = 1$. □

例 5 在由所有实多项式组成的线性空间中, 定义

$$(f, g) = \int_0^\infty \mathrm{e}^{-t}f(t)g(t)\mathrm{d}t.$$

由于被积函数包含指数因子, 上式中的广义积分对所有多项式 f 和 g 都收敛. □

定理 3.8 在欧氏空间 V 中, 任意内积都满足 Cauchy-Schwarz 不等式: 对 V 中任意 x 和 y, 都有

$$|(x, y)|^2 \leqslant (x, x)(y, y).$$

更进一步, 当且仅当 x 和 y 相关时, 上式中的等号成立.

证明 若 $x = O$ 或 $y = O$, 那么结论显然成立, 所以我们可以假设 x 和 y 都不为零. 令 $z = ax + by$, 其中 a 和 b 为待定纯量. 对所有 a 和 b, 有不等式 $(z, z) \geqslant 0$. 当我们用 x 和 y 表示该不等式且适当地选取 a 和 b 时, 将得到 Cauchy-Schwarz 不

等式.

为了用 x 和 y 表示 (z,z), 由性质 $(1')$, (2), $(3')$, 得

$$\begin{aligned}(z,z) &= (ax+by, ax+by) = (ax,ax)+(ax,by)+(by,ax)+(by,by)\\ &= a\bar{a}(x,x)+a\bar{b}(x,y)+b\bar{a}(y,x)+b\bar{b}(y,y) \geqslant 0.\end{aligned}$$

令 $a=(y,y)$ 并在上式中消去正因子 (y,y) 可得

$$(y,y)(x,x)+\bar{b}(x,y)+b(y,x)+b\bar{b} \geqslant 0.$$

再令 $b=-(x,y)$, 则 $\bar{b}=-(y,x)$ 且上式简化为

$$(y,y)(x,x) \geqslant (x,y)(y,x).$$

由于 $(x,y)(y,x)=|(x,y)|^2$, 因此我们证明了 Cauchy-Schwarz 不等式成立. 从证明的过程中可知当且仅当 $z=O$ 时, 即当且仅当 x 和 y 线性相关时, 不等式中的等号成立. □

例 当我们对例 3 中的实欧氏空间应用 Cauchy-Schwarz 不等式时, 我们将得到下述积分不等式

$$\left(\int_a^b f(t)g(t)\mathrm{d}t\right)^2 \leqslant \left(\int_a^b f^2(t)\mathrm{d}t\right)\left(\int_a^b g^2(t)\mathrm{d}t\right). \qquad \square$$

我们用内积为欧氏空间引入一个长度的度量概念.

定义 *在欧氏空间 V 中, 称非负数 $(x,x)^{1/2}$ 为元素 x 的***范数** (norm), *记作 $\|x\|$.*

若我们用范数表示 Cauchy-Schwarz 不等式, 它可写成如下形式:

$$|(x,y)| \leqslant \|x\|\|y\|.$$

由于可在同一空间中定义多个不同的内积, 元素的范数依赖于所用的内积, 这种不确定性是可以理解的. 这与我们可按不同的长度单位用不同的数来表示同一个给定线段的长度类似. 下述定理给出范数的基本性质, 这些性质与内积的选取方式无关.

定理 3.9 *在欧氏空间中, 对任意元素 x 和 y 以及任意纯量 c, 所有范数都满足下述各性质:*

(a) $\|x\|=0$, 若 $x=O$.

(b) $\|x\|>0$, 若 $x\neq O$ (正次定性).

(c) $\|cx\|=|c|\|x\|$ (齐次性).

(d) $\|x+y\| \leqslant \|x\|+\|y\|$ (三角不等式).

当 $x=O$, 或 $y=O$, 或存在 $c>0$ 使得 $y=cx$ 时, (d) 中的等号成立.

证明 由内积公理易证 (a), (b), (c), 接下来我们证明 (d). 注意到

$$\begin{aligned}\|x+y\|^2 &= (x+y,x+y) = (x,x)+(y,y)+(x,y)+(y,x)\\ &= \|x\|^2+\|y\|^2+(x,y)+\overline{(x,y)}.\end{aligned}$$

和 $(x,y)+\overline{(x,y)}$ 是实数. 由 Cauchy-Schwarz 不等式, 我们有 $|(x,y)|\leqslant\|x\|\|y\|$ 以及 $|\overline{(x,y)}|\leqslant\|x\|\|y\|$, 所以

$$\|x+y\|^2\leqslant\|x\|^2+\|y\|^2+2\|x\|\|y\|=(\|x\|+\|y\|)^2$$

因此 (d) 得证. 当 $y=cx$, 其中 $c>0$ 时, 我们有

$$\|x+y\|=\|x+cx\|=(1+c)\|x\|=\|x\|+\|cx\|=\|x\|+\|y\|. \qquad \square$$

定义 在实欧氏空间 V 中, 定义两个非零元素 x 和 y 的夹角为满足 $0\leqslant\theta\leqslant\pi$ 以及等式

$$\cos\theta=\frac{(x,y)}{\|x\|\|y\|} \tag{3.8}$$

的数 θ. □

注 Cauchy-Schwarz 不等式表明等式 (3.8) 中等号右边的值落在区间 $[-1,1]$ 中, 所以在 $[0,\pi]$ 中恰有一个 θ 满足等式 (3.8).

3.12 欧氏空间中的正交性

定义 设 x 和 y 为欧氏空间 V 中的两个元素, 如果它们的内积为零, 则称它们正交. 若对 V 的子集 S 中的任意两个不同元素 x 和 y 都有 $(x,y)=0$, 则称 S 为**正交组** (orthogonal set). 如果一个正交组中的每一个元素的范数都等于 1, 则称它为**标准正交组** (orthonormal set).

元素 O 与 V 中所有元素都正交, 而且 (由公理 4) 它是唯一一个与自己正交的元素.

我们知道对 $\mathbb{R}^n$ 中两个向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 当且仅当它们满足 Pythagoras 恒等式 $\|\boldsymbol{A}+\boldsymbol{B}\|^2=\|\boldsymbol{A}\|^2+\|\boldsymbol{B}\|^2$ 时, 它们正交. 下面的定理将这个结果推广到所有欧氏空间中.

定理 3.10 对欧氏空间 V 中的任意两个元素 x 和 y, 当且仅当它们满足 Pythagoras 恒等式

$$\|x+y\|^2=\|x\|^2+\|y\|^2.$$

时, 它们正交.

证明 由公式

$$\|x+y\|^2=(x+y,x+y)=\|x\|^2+\|y\|^2+(x,y)+(y,x)$$

可立即推出本定理. □

接下来我们研究正交性与线性无关性之间的关系.

定理 3.11 在欧氏空间 V 中, 任意一个不包含零元素的正交组都线性无关. 特别地, 在一个 n 维欧氏空间中, 任意一个包含 n 个非零元素的正交组都是 V 的基.

证明 令 S 为 V 中一个不包含零元素的正交组, 考虑任意表示零元素的 S 中元素的有限线性组合, 设这个组合为

$$\sum_{i=1}^{k} c_i x_i = O,$$

其中所有 $x_i \in S$. 取 x_1 与上式两边的内积, 由于对 $i \neq 1$, 都有 $(x_1, x_i) = 0$, 所以我们有 $c_1(x_1, x_1) = 0$. 又因为 $x_1 \neq O$, 所以 $(x_1, x_1) \neq 0$, 因此 $c_1 = 0$. 用 x_j 取代 x_1 重复上述过程, 即得所有 $c_j = 0$. 这证明了 S 线性无关. 如果 $\dim V = n$ 且 S 包含 n 个元素, 那么由定理 3.7 (b) 可知 S 是 V 的基. □

例 (要求读者了解积分.) 设在实线性空间 $C(0, 2\pi)$ 中的内积定义为 $(f, g) = \int_0^{2\pi} f(x)g(x)\mathrm{d}x$, 令 S 表示下述三角函数组成的集合 $\{u_0, u_1, u_2, \cdots\}$:

$$u_0(x) = 1, \quad u_{2n-1}(x) = \cos nx, \quad u_{2n}(x) = \sin nx, \quad n = 1, 2, \cdots.$$

已知对 $m \neq n$, 都有

$$\int_0^{2\pi} u_n(x) u_m(x)\mathrm{d}x = 0.$$

(这称为 $\cos nx$ 与 $\sin nx$ 的正交性.) 因此, S 是正交组. 又因为 S 中元素都不为零, 所以 S 线性无关.

求 S 中所有元素的范数十分简单. 考虑到 $(u_0, u_0) = \int_0^{2\pi} \mathrm{d}x = 2\pi$, 以及对 $n \geqslant 1$ 有

$$(u_{2n-1}, u_{2n-1}) = \int_0^{2\pi} \cos^2 nx\mathrm{d}x = \pi, \quad (u_{2n}, u_{2n}) = \int_0^{2\pi} \sin^2 nx\mathrm{d}x = \pi.$$

于是 $\|u_0\| = \sqrt{2\pi}$, 且对 $n \geqslant 1$, 都有 $\|u_n\| = \sqrt{\pi}$. 对每一个 u_n, 用它的范数除它本身可得标准正交组 $\{\phi_0, \phi_1, \phi_2, \cdots\}$, 其中 $\phi_n = u_n/\|u_n\|$. 因此对 $n \geqslant 1$, 我们有

$$\varphi_0(x) = \frac{1}{\sqrt{2\pi}}, \quad \varphi_{2n-1}(x) = \frac{\cos nx}{\sqrt{\pi}}, \quad \varphi_{2n}(x) = \frac{\sin nx}{\sqrt{\pi}}.$$ □

在 3.14 节中我们将证明所有有限维欧氏空间都有正交基. 下面的定理给出求一个元素关于正交基的各分量的方法.

定理 3.12 令 V 为一个 n 维欧氏空间, 设 $S = \{e_1, \cdots, e_n\}$ 是 V 的正交基, 当 V 中元素表示成基 S 中元素的线性组合, 比如

$$x = \sum_{i=1}^{k} c_i e_i, \tag{3.9}$$

时, x 关于有序基 $(e_1, \cdots, e_n)$ 的各分量由公式

$$c_j = \frac{(x, e_j)}{(e_j, e_j)}, \quad j = 1, 2, \cdots, n. \tag{3.10}$$

给出, 特别当 S 为标准正交基时, c_j 由公式

$$c_j = (x, e_j) \tag{3.11}$$

给出.

证明 取 e_j 与展开式 (3.9) 两边的内积, 因为当 $i \neq j$ 时有 $(e_i, e_j) = 0$, 所以我们有

$$(x, e_j) = \sum_{i=1}^{k} c_i(e_i, e_j) = c_j(e_j, e_j)$$

由此可得公式 (3.10) 且当 $(e_i, e_j) = 1$ 时可得公式 (3.11). □

若 $\{e_1, \cdots, e_n\}$ 为标准正交基, 那么展开式 (3.9) 可写成如下形式:

$$x = \sum_{i=1}^{n}(x, e_i)e_i \tag{3.12}$$

下一条定理表明, 在一个给定了一组标准正交基的有限维欧氏空间中, 可通过一个类似于 $\mathbb{R}^n$ 中向量点积公式的公式用分量来求两个元素的内积.

定理 3.13 令 V 为一个 n 维欧氏空间, 设 $\{e_1, \cdots, e_n\}$ 是 V 的一组标准正交基, 那么 V 中任意两个元素 x 和 y 的内积可由

$$(x, y) = \sum_{i=1}^{n}(x, e_i)\overline{(y, e_i)} \tag{3.13}$$

确定.

特别地, 当 $x = y$ 时, 我们有

$$\|x\|^2 = \sum_{i=1}^{n}|(x, e_i)|^2 \tag{3.14}$$

证明 取 y 与表达式 (3.12) 两边的点积并由内积的线性性即得公式 (3.13). 当 $x = y$ 时, 公式 (3.13) 化为公式 (3.14). □

注 为了纪念 M. A. Parseval(约 1776 — 1836), 我们称公式 (3.13) 为 Parseval 公式, 他为一个特殊函数空间推导了这样一个公式. 等式 (3.14) 是 Parseval 公式的特殊情形, 它可看作是 Pythagoras 定理的推广.

3.13 习 题

1. 令 $\boldsymbol{x} = (x_1, \cdots, x_n)$ 和 $\boldsymbol{y} = (y_1, \cdots, y_n)$ 为 $\mathbb{R}^n$ 中的任意向量, 若 $(\boldsymbol{x}, \boldsymbol{y})$ 由下列各小题中给出的公式所定义, 判断 $(\boldsymbol{x}, \boldsymbol{y})$ 是否是 $\mathbb{R}^n$ 的内积, 如果不是, 指出它违反了哪条公理.

   (a) $(\boldsymbol{x}, \boldsymbol{y}) = \sum\limits_{i=1}^{n} x_i|y_i|$.

   (b) $(\boldsymbol{x}, \boldsymbol{y}) = \left|\sum\limits_{i=1}^{n} x_i y_i\right|$.

   (c) $(\boldsymbol{x}, \boldsymbol{y}) = \sum\limits_{i=1}^{n} x_i \sum\limits_{j=1}^{n} y_j$.

   (d) $(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum\limits_{i=1}^{n} x_i^2 y_i^2\right)^{1/2}$.

   (e) $(\boldsymbol{x}, \boldsymbol{y}) = \sum\limits_{i=1}^{n} (x_i + y_i)^2 - \sum\limits_{i=1}^{n} x_i^2 - \sum\limits_{i=1}^{n} y_i^2$.

2. 设我们保留实内积的前三条公理 (对称性、线性性和齐次性), 但公理 4 则用下述新公理 (公理 4′) 代替: 当且仅当 $x = O$ 时 $(x,x)=0$. 证明: 要么对所有 $x \neq O$, 都有 $(x,x)>0$, 要么对所有 $x \neq O$, 都有 $(x,x)<0$.

   [提示: 设存在 x, y, 使得 $(x,x)>0$, $(y,y)<0$, 在 $\{x,y\}$ 生成的空间中找出一个元素 $z \neq O$, 但是有 $(z,z)=0$.]

   对实欧氏空间中的任意元素 x 和 y, 证明习题 3~7 题中的各命题都成立.

3. 当且仅当 $\|x+y\| = \|x-y\|$ 时 $(x,y)=0$.
4. 当且仅当 $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ 时 $(x,y)=0$.
5. 当且仅当对任意实数 c 都有 $\|x+cy\| \geqslant \|x\|$ 时 $(x,y)=0$.
6. 当且仅当 $\|x\| = \|y\|$ 时 $(x+y, x-y)=0$.
7. 若 x 和 y 为非零元素且它们的夹角为 θ, 那么

$$\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos\theta.$$

8. 证明下列各恒等式在所有欧氏空间中都成立.
   (a) $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + (x,y) + (y,x)$.
   (b) $\|x+y\|^2 - \|x-y\|^2 = 2(x,y) + 2(y,x)$.
   (c) $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$.
9. 在复欧氏空间中, 证明: 对任意元素 x, y, z 和任意复数 a, b, 内积都有下述性质.
   (a) $(ax, by) = a\bar{b}(x,y)$.　　　(b) $(x, ay+bz) = \bar{a}(x,y) + \bar{b}(x,z)$.
10. 在由次数 $\leqslant n$ 的所有多项式组成的线性空间 P_n 中定义

$$(f,g) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) g\left(\frac{k}{n}\right).$$

   (a) 证明 (f,g) 是 P_n 的内积.
   (b) 当 $f(t)=t$, $g(t)=at+b$ 时, 求 (f,g).
   (c) 若 $f(t)=t$, 求正交于 f 的所有线性多项式 g.

   习题 11~17 用到微积分知识.

11. 在实线性空间 $C(1,\mathrm{e})$ 中, 由公式

$$(f,g) = \int_1^{\mathrm{e}} (\ln x) f(x) g(x) \mathrm{d}x$$

   定义内积.
   (a) 当 $f(x)=\sqrt{x}$ 时, 计算 $\|f\|$ 的值.
   (b) 求与常数函数 $f(x)=1$ 正交的一个线性多项式 $g(x)=a+bx$.
12. 在实线性空间 $C(-1,1)$ 中, 令 $(f,g)=\int_{-1}^{1} f(t)g(t)\mathrm{d}t$, 考虑以下三个函数

$$u_1(t)=1, \quad u_2(t)=t, \quad u_3(t)=1+t.$$

   证明它们中有两个互相正交, 有两个夹角为 $\pi/3$, 有两个夹角为 $\pi/6$.
13. 在所有实多项式组成的线性空间中, 定义 $(f,g)=\int_0^{\infty} \mathrm{e}^{-t} f(t)g(t)\mathrm{d}t$.
   (a) 证明对任意多项式 f 和 g, 这个广义积分都绝对收敛.
   (b) 若对 $n=0, 1, 2, \cdots$, 都有 $x_n(t)=t^n$, 证明 $(x_n, x_m)=(m+n)!$.
   (c) 当 $f(t)=(t+1)^2$ 且 $g(t)=t^2+1$ 时, 计算 (f,g).
   (d) 求与 $f(t)=1+t$ 正交的所有线性多项式 $g(t)=a+bt$.
14. 在所有实多项式组成的线性空间中, 当 (f,g) 由以下各公式定义时, 判断 (f,g) 是否为内积, 如果 (f,g) 不是内积, 指出它违反了哪条公理. 在 (c) 中, f' 和 g' 分别表示 f 和 g 的导数.
   (a) $(f,g)=f(1)g(1)$.　　　(b) $(f,g)=\left|\int_0^1 f(t)g(t)\mathrm{d}t\right|$.
   (c) $(f,g)=\int_0^1 f'(t)g'(t)\mathrm{d}t$.　　　(d) $(f,g)=\left(\int_0^1 f(t)\mathrm{d}t\right)\left(\int_0^1 g(t)\mathrm{d}t\right)$.

15. 令 V 是在 $[0,+\infty)$ 上连续且积分 $\int_0^{\infty} e^{-t}f^2(t)dt$ 收敛的所有实函数 f 组成的空间, 定义 $(f,g)=\int_0^{\infty} e^{-t}f(t)g(t)dt$,

 (a) 证明对 V 中任意 f 和 g, (f,g) 中的积分都绝对收敛.

 [提示: 用 Cauchy-Schwarz 不等式估算积分 $\int_0^{M} e^{-t}|f(t)g(t)|dt$.]

 (b) 证明 V 是一个以 (f,g) 为内积的线性空间.

 (c) 对于 $f(t)=e^{-t}$ 和 $g(t)=t^n$, 其中 $n=0,1,2,\cdots$, 计算 (f,g).

16. 设 V 包含满足级数 $\sum\limits_{n=1}^{\infty} x_n^2$ 收敛的所有实数序列 $\{x_n\}$, 对 V 中两个元素 $x=\{x_n\}$ 和 $y=\{y_n\}$ 定义

$$(x,y)=\sum_{n=1}^{\infty} x_n y_n$$

 (a) 证明这个级数绝对收敛.

 $\left[\text{提示: 用 Cauchy-Schwarz 不等式估算和 } \sum\limits_{n=1}^{M} |x_n y_n|\right]$.

 (b) 证明 V 是一个以 (x,y) 为内积的线性空间.

 (c) 当对 $n\geqslant 1$, 都有 $x_n=1/n$ 以及 $y_n=1/(n+1)$ 时, 求 (x,y).

 (d) 当对 $n\geqslant 1$, 都有 $x_n=2^{-n}$ 以及 $y_n=1/n!$ 时, 求 (x,y).

17. 在区间 $[a,b]$ 上连续的所有复值函数组成的空间中, 对给定的在 $[a,b]$ 上连续的正函数 w, 若我们用公式

$$(f,g)=\int_a^b w(t)f(t)\overline{g(t)}dt.$$

 定义内积, 证明由此得到的是一个复欧氏空间.

3.14 正交组的构造 · Gram-Schmidt 方法

任意有限维线性空间都有基, 如果这个空间是欧氏空间, 那么我们总可以构造一组正交基. 这个结果可作为一个一般性定理的推论而导出. 这个一般性定理的证明将会告诉我们如何在任意一个欧氏空间中构造一组正交基, 而不管这个空间是有限维的还是无限维的. 为了纪念 J. P. Gram(1850 — 1916) 和 E. Schmidt(1845 — 1921), 我们将这种构造法称为**Gram-Schmidt 正交化方法**(Gram-Schmidt orthogonalization process).

定理 3.14(正交化定理) *设 $x_1,x_2,\cdots$ 为欧氏空间 V 中元素的序列 (它可以是有限的也可以是无限的), 令 $L(x_1,\cdots,x_k)$ 表示该序列中前 k 个元素生成的子空间, 那么相应地存在 V 中元素的序列 $y_1,y_2,\cdots$, 使得对任意正整数 k, 它都具有下述各性质.*

(a) 元素 y_k 与子空间 $L(y_1,\cdots,y_{k-1})$ 中的所有元素都正交.

(b) $y_1,\cdots,y_k$ 生成的子空间与 $x_1,\cdots,x_k$ 生成的子空间相同, 即

$$L(y_1,\cdots,y_k)=L(x_1,\cdots,x_k).$$

(c) 序列 $y_1,y_2,\cdots$ 在允许相差一个纯量因子的意义下唯一. 即若 $z_1,z_2,\cdots$ 是 V 中元素的序列且对任意 k 它都满足性质 (a) 和 (b), 那么对任意 k, 都存在纯量

c_k 使得 $z_k = c_k y_k$.

证明　我们归纳地构造元素 $y_1, y_2, \cdots$. 首先, 我们令 $y_1 = x_1$, 接下来我们假设已经构造了 $y_1, \cdots, y_r$, 且对 $k \leqslant r$, 它们满足性质 (a) 和 (b), 此时我们令

$$y_{r+1} = x_{r+1} - \sum_{i=1}^{r} a_i y_i, \tag{3.15}$$

其中各纯量 a_i 依下述方式选取. 由于当 $i \neq j$ 时, $(y_i, y_j) = 0$, 所以 y_{r+1} 与 y_j 的内积由

$$(y_{r+1}, y_j) = (x_{r+1}, y_j) - \sum_{i=1}^{r} a_i (y_i, y_j) = (x_{r+1}, y_j) - a_j (y_j, y_j)$$

给出. 若 $y_j \neq O$, 我们可令

$$a_j = \frac{(x_{r+1}, y_j)}{(y_j, y_j)}. \tag{3.16}$$

从而使得 y_{r+1} 与 y_j 正交; 若 $y_j = O$, 则可任意选 a_j, 此时 y_{r+1} 总与 y_j 正交, 这里我们令 $a_j = 0$. 这样, 元素 y_{r+1} 就完全确定了并且它与前面的所有元素 $y_1, \cdots, y_r$ 都正交, 所以它与子空间

$$L(y_1, \cdots, y_r)$$

中所有元素正交, 故当 $k = r + 1$ 时, (a) 得证.

为了在 $k = r + 1$ 时证明 (b), 我们必须证明: 若 $L(y_1, \cdots, y_r) = L(x_1, \cdots, x_r)$, 那么 $L(y_1, \cdots, y_{r+1}) = L(x_1, \cdots, x_{r+1})$. 显然序列中的前 r 个元素 $y_1, \cdots, y_r$ 属于 $L(x_1, \cdots, x_r)$, 所以它们也属于更大的子空间 $L(x_1, \cdots, x_{r+1})$. 又因为由等式 (3.15) 定义的新元素 y_{r+1} 为 $L(x_1, \cdots, x_{r+1})$ 中两个元素之差, 所以它也属于 $L(x_1, \cdots, x_{r+1})$. 这就证明了

$$L(y_1, \cdots, y_{r+1}) \subseteq L(x_1, \cdots, x_{r+1}).$$

等式 (3.15) 表明 x_{r+1} 是 $L(y_1, \cdots, y_{r+1})$ 中两个元素之和, 因此同理可知

$$L(x_1, \cdots, x_{r+1}) \subseteq L(y_1, \cdots, y_{r+1}).$$

所以当 $k = r + 1$ 时, (b) 得证. 于是通过对 k 用归纳法, 我们证明了 (a) 和 (b).

最后我们对 k 用归纳法来证明 (c). 在 $k = 1$ 的情形下结论显然成立, 因此我们假设当 $k = r$ 时 (c) 为真. 下面考虑元素 z_{r+1}. 由 (b) 可知 z_{r+1} 属于子空间

$$L(y_1, \cdots, y_{r+1}),$$

所以我们可将它写成

$$z_{r+1} = \sum_{i=1}^{r+1} c_i y_i = v_r + c_{r+1} y_{r+1},$$

其中 $v_r \in L(y_1, \cdots, y_r)$. 我们希望证明 $v_r = O$. 由性质 (a), z_{r+1} 和 $c_{r+1} y_{r+1}$ 都与 v_r 正交, 所以它们的差 v_r 也与 v_r 正交, 即 v_r 与自己正交, 所以 $v_r = O$, 从而正交化定理得证. □

在上述构造中，假定对某个 r 有 $y_{r+1}=O$，则等式 (3.15) 表明 x_{r+1} 是 $y_1,\cdots,y_r$ 的线性组合，所以它也是 $x_1,\cdots,x_r$ 的线性组合，因此 $x_1,\cdots,x_{r+1}$ 线性相关．也就是说，如果前 k 个元素 $x_1,\cdots,x_k$ 线性无关，那么相应的元素 $y_1,\cdots,y_k$ 都是非零元素．此时等式 (3.15) 中的系数 a_i 由公式 (3.16) 确定，定义 $y_1,\cdots,y_k$ 的公式转化为，对 $r=1,2,\cdots,k-1$，

$$y_1=x_1,\quad y_{r+1}=x_{r+1}-\sum_{i=1}^{r}\frac{(x_{r+1},y_i)}{(y_i,y_i)}y_i. \tag{3.17}$$

这些公式给出了构造包含非零元素 $y_1,\cdots,y_k$ 的正交组，使得它所生成的子空间与给定无关组 $x_1,\cdots,x_k$ 所生成的子空间相同的 Gram-Schmidt 方法．特别地，当 $x_1,\cdots,x_k$ 是一个有限维欧氏空间的基时，$y_1,\cdots,y_k$ 是该空间的正交基．我们可进一步将正交组中每个元素 y_j 单位化从而得到它的标准正交基．为此，只需用一个元素的范数除该元素本身即可．因此，作为定理 3.14 的推论，我们可得如下定理．

定理 3.15 *所有有限维欧氏空间都有标准正交基．* □

定义 *设 x 和 y 为某欧氏空间中的元素且 $y\neq O$，那么元素*

$$\frac{(x,y)}{(y,y)}y$$

*称为 x 在 y 上的***投影***．*

图 1.11 给出了投影在 $\mathbb{R}^2$ 中的几何意义．在应用 Gram-Schmidt 方法 (公式 (3.17)) 的过程中，我们通过用 x_{r+1} 减去 x_{r+1} 在已经求得的各元素 $y_1,\cdots,y_r$ 上的投影来求得新元素 y_{r+1}．图 3.1 给出了在向量空间 $\mathbb{R}^3$ 中这种构造法的几何表示．

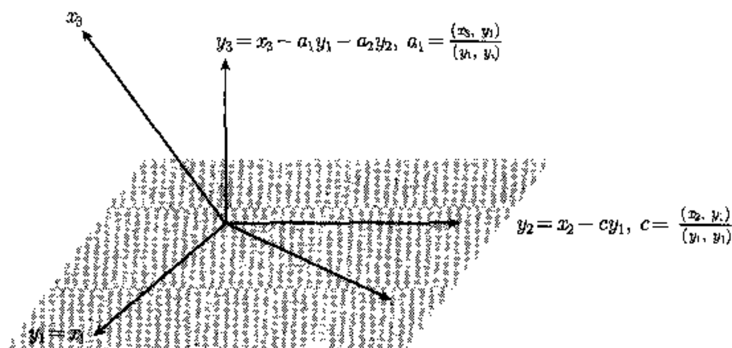


图 3.1 $\mathbb{R}^3$ 中的 Gram-Schmidt 方法．由给定线性无关组 $\{x_1,x_2,x_3\}$ 构造的正交组 $\{y_1,y_2,y_3\}$

例 1 在 $\mathbb{R}^4$ 中，求由三个向量 $\boldsymbol{x}_1=(1,-1,1,-1)$，$\boldsymbol{x}_2=(5,1,1,1)$，$\boldsymbol{x}_3=(-3,-3,1,-3)$ 生成的子空间的一组标准正交基．

解 应用 Gram-Schmidt 方法我们可得

$$y_1 = x_1 = (1, -1, 1, -1),$$

$$y_2 = x_2 - \frac{(x_2, y_1)}{(y_1, y_1)} y_1 = x_2 - y_1 = (4, 2, 0, 2),$$

$$y_3 = x_3 - \frac{(x_3, y_1)}{(y_1, y_1)} y_1 - \frac{(x_3, y_2)}{(y_2, y_2)} y_2 = x_2 - y_1 + y_2 = (0, 0, 0, 0).$$

由于 $y_3 = O$, 向量 x_1, x_2, x_3 必相关, 又由于 y_1 和 y_2 都不是零向量, 所以向量 x_1 和 x_2 线性无关, 因此 $L(x_1, x_2, x_3)$ 是一个二维子空间. 集合 $\{y_1, y_2\}$ 是它的一组正交基. 用 y_1 和 y_2 的范数分别除它们本身所得到的两个向量

$$\frac{y_1}{\|y_1\|} = \frac{1}{2}(1, -1, 1, -1), \quad \frac{y_2}{\|y_2\|} = \frac{1}{\sqrt{6}}(2, 1, 0, 1).$$

组成该空间的一组标准正交基. □

例 2 (Legendre 多项式) (要求读者熟悉微积分.) 在由多项式组成的线性空间中, 令内积为 $(x, y) = \int_{-1}^{1} x(t)y(t)\mathrm{d}t$. 考虑无穷序列 $x_0, x_1, x_2, \cdots$, 其中 $x_n(t) = t^n$. 对该序列应用正交化定理后, 我们将得到一个新的多项式序列 $y_0, y_1, y_2, \cdots$, 这个序列由法国数学家 A. M. Legendre (1752—1833) 在研究势论的过程中首先发现. 通过 Gram-Schmidt 方法易求该序列的前几个多项式. 首先, 我们知道 $y_0(t) = x_0(t) = 1$. 由于

$$(y_0, y_0) = \int_{-1}^{1} \mathrm{d}t = 2, \quad (x_1, y_0) = \int_{-1}^{1} t\mathrm{d}t = 0,$$

可得

$$y_1(t) = x_1(t) - \frac{(x_1, y_0)}{(y_0, y_0)} y_0(t) = x_1(t) = t.$$

接下来, 由关系

$$(x_2, y_0) = \int_{-1}^{1} t^2 \mathrm{d}t = \frac{2}{3}, \quad (x_2, y_1) = \int_{-1}^{1} t^3 \mathrm{d}t = 0, \quad (y_1, y_1) = \int_{-1}^{1} t^2 \mathrm{d}t = \frac{2}{3},$$

可得

$$y_2(t) = x_2(t) - \frac{(x_2, y_0)}{(y_0, y_0)} y_0(t) - \frac{(x_2, y_1)}{(y_1, y_1)} y_1(t) = t^2 - \frac{1}{3}.$$

同理可得

$$y_3(t) = t^3 - \frac{3}{5}t, \quad y_4(t) = t^4 - \frac{6}{7}t^2 + \frac{3}{35}, \quad y_5(t) = t^5 - \frac{10}{9}t^3 + \frac{5}{21}t.$$

在研究微分方程时我们需要用到这些多项式, 其中已经证明了

$$y_n(t) = \frac{n!}{(2n)!} \frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2 - 1)^n.$$

由

$$P_n(t) = \frac{(2n)!}{2^n (n!)^2} y_n(t) = \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}t^n}(t^2 - 1)^n$$

确定的多项式 P_n 称为 Lengendre 多项式. 令 $\varphi_n = y_n / \|y_n\|$, 则对应的标准正交序列 $\varphi_0, \varphi_1, \varphi_2, \cdots$ 中的多项式称为标准 Legendre 多项式. 由于 $y_0, \cdots, y_5$ 由前述公

式确定, 我们可求得前几个标准 Legendre 多项式为

$$\varphi_0(t)=\sqrt{\frac{1}{2}},\quad \varphi_1(t)=\sqrt{\frac{3}{2}}t,\quad \varphi_2(t)=\frac{1}{2}\sqrt{\frac{5}{2}}(3t^2-1),\quad \varphi_3(t)=\frac{1}{2}\sqrt{\frac{7}{2}}(5t^3-3t),$$

$$\varphi_4(t)=\frac{1}{8}\sqrt{\frac{9}{2}}(35t^4-30t^2+3),\quad \varphi_5(t)=\frac{1}{8}\sqrt{\frac{11}{2}}(63t^5-70t^3+15t).$$ □

3.15 正交补 · 投影

设 V 是一个欧氏空间, 令 S 为 V 的一个有限维子空间. 现在讨论如下逼近问题: *给定 V 中元素 x, 求 S 中元素, 使得它到 x 的距离最小*. 对 V 中两个元素 x 和 y, 范数 $\|x-y\|$ 称为 x 与 y 之间的距离.

在对一般情形讨论此问题之前, 我们先考虑如图 3.2 所示的特殊情形. 此时 $V=\mathbb{R}^3$, S 是它的二维子空间, 即 S 是过原点的平面. 给定 V 中元素 x, 那么该问题转化为在平面 S 中求离 x 最近的点 s.

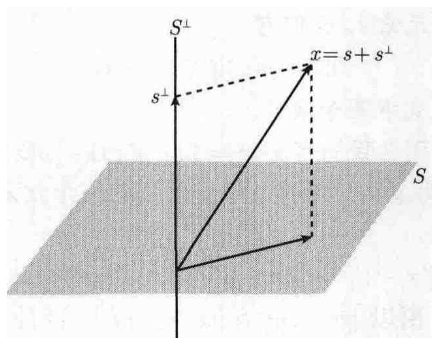


图 3.2 $\mathbb{R}^3$ 中过原点的平面的正交补是一条垂直于该平面的直线.
图中 $s=p(x)$ 且 $S^\perp=x-p(x)$

如果 $x\in S$, 显然 $s=x$ 是该问题的解. 如果 $x\notin S$, 那么可过点 x 作 S 的垂线, 从而求得距 x 最近的点. 这个简单的例子给出了求解一般逼近问题的一种方法并引出下述讨论.

定义 *设 S 为欧氏空间 V 的子集, 若 V 中一个元素与 S 中的所有元素都正交, 那么我们称该元素与 S 正交. 称所有与 S 正交的元素组成的集合为"S 垂线"并记作 $S^\perp$.*

易证无论 S 是否为 V 的子空间, $S^\perp$ 都是 V 的子空间. 当 S 也是子空间时, 称 $S^\perp$ 为 S 的**正交补** (orthogonal complement).

例 如图 3.2 所示, 如果 S 是过原点的一个平面, 那么 $S^\perp$ 是过原点垂直于该平面的直线. □

定义　对任意欧氏空间 V, 令 S 为它的一个有限维子空间且 $\{e_1,\cdots,e_n\}$ 是 S 的标准正交基. 若 $x\in V$, 那么由方程

$$p(x)\sum_{i=1}^{n}(x,e_i)e_i \tag{3.18}$$

定义的 S 中元素 $p(x)$ 称为 x 在子空间 S 上的投影.

注意到 $(p(x),e_k)=(x,e_k)$, 因此对基中每个元素 e_k, $(x-p(x),e_k)=0$, 所以 $x-p(x)\in S^{\perp}$. 在下一节中我们将证明投影 $p(x)$ 是本节开始提到的逼近问题的解.

3.16　用有限维子空间中的元素给出欧氏空间中元素的最优逼近

定理 3.16 (**逼近定理**)　设 S 为欧氏空间 V 的一个有限维子空间, 令 x 为 V 中任意元素, 那么 x 在 S 上的投影 $p(x)$ 与 x 的距离比 S 上其他任意点与 x 的距离都小. 即对 S 中所有元素 t, 我们有

$$\|x-p(x)\|\leqslant\|x-t\|$$

当且仅当 $t=p(x)$ 时上式中等号成立.

证明　如果 $t\in S$, 那么我们有 $x-t=\{x-p(x)\}+\{p(x)-t\}$. 因为 $x-p(x)\in S^{\perp}$, $p(x)-t\in S$, 所以我们将 $x-t$ 表示为相互正交的两个元素之和, 由 Pythagoras 恒等, 我们有

$$\|x-t\|^2=\|x-p(x)\|^2+\|p(x)-t\|^2$$

又因为 $\|p(x)-t\|^2\geqslant 0$, 所以 $\|x-t\|^2\geqslant\|x-p(x)\|^2$, 当且仅当 $t=p(x)$ 时, 等号成立. □

注　因为 x 是相互垂直的两个元素 $p(x)$ 和 $x-p(x)$ 之和, 所以由 Pythagoras 恒等式可知 $\|x\|^2=\|p(x)\|^2+\|x-p(x)\|^2$, 所以从 x 到子空间 S 的最短距离可由关系

$$\|x-p(x)\|^2=\|x\|^2-\|p(x)\|^2 \tag{3.19}$$

求得.

计算 $\|p(x)\|$ 最简单的方法是用 $p(x)$ 的定义 (公式 (3.18)), 此时

$$\|p(x)\|^2=\sum_{i=1}^{n}|(x,e_i)|^2.$$

下面两个用到微积分的例子演示了逼近定理的用法.

例 1 (用三角多项式逼近 $[0,2\pi]$ 上的连续函数)　设 $V=C(0,2\pi)$ 是在区间 $[0,2\pi]$ 上连续的所有实函数的线性空间, 定义它的内积为通常意义下的积分 $(f,g)=\int_0^{2\pi}f(x)g(x)\mathrm{d}x$. 在 3.12 节中, 我们已经求得了由三角函数 $\varphi_0,\varphi_1,\varphi_2,\cdots$ 组成的 V 的一组标准正交基, 其中

$$\varphi_0(x)=\frac{1}{\sqrt{2\pi}},\quad \varphi_{2k-1}(x)=\frac{\cos kx}{\sqrt{\pi}},\quad \varphi_{2k}(x)=\frac{\sin kx}{\sqrt{\pi}},\quad k\geqslant 1. \tag{3.20}$$

$2n+1$ 个元素 $\varphi_0,\varphi_1,\cdots,\varphi_{2n}$ 生成 V 的一个 $2n+1$ 维子空间 S, 我们称 S 中的元素为**三角多项式** (trigonometric polynomial)

如果 $f\in C(0,2\pi)$, 令 f_n 表示 f 在子空间 S 上的投影, 那么我们有

$$f_n=\sum_{k=0}^{2n}(f,\varphi_k)\varphi_k,\quad 其中\quad (f,\varphi_k)\int_0^{2\pi}f(x)\varphi_k(x)\mathrm{d}x. \tag{3.21}$$

我们称数 (f,φ_k) 为 f 的 **Fourier 系数**. 将公式 (3.20) 代入表达式 (3.21) 可得

$$f_n(x)=\frac{1}{2}a_0+\sum_{k=1}^{n}(a_k\cos kx+b_k\sin kx), \tag{3.22}$$

其中对 $k=0,1,2,\cdots,n$,

$$a_k=\frac{1}{\pi}\int_0^{2\pi}f(x)\cos kx\mathrm{d}x,\quad b_k=\frac{1}{\pi}\int_0^{2\pi}f(x)\sin kx\mathrm{d}x,$$

逼近定理告诉我们, 在使范数 $\|f-f_n\|$ 越小越好的意义下, 用公式 (3.22) 中的三角多项式逼近 f 比用 S 中任何其他三角多项式逼近都要好. □

例 2 (用次数 $\leqslant n$ 的多项式逼近区间 $[-1,1]$ 上的连续函数) 设 $V=C(-1,1)$ 是在 $[-1,1]$ 上连续的所有实函数组成的空间, 用等式 $(f,g)=\int_{-1}^{1}f(x)g(x)\mathrm{d}x$ 定义该空间的内积, 那么 3.14 节中引入的$n+1$个标准 Legendre 多项式$\varphi_0,\varphi_1,\cdots,\varphi_n$ 生成 V 的一个由所有次数 $\leqslant n$ 的多项式组成的 $n+1$ 维子空间 S. 如果 $f\in C(-1,1)$, 令 f_n 表示 f 在 S 上的投影, 那么我们有

$$f_n=\sum_{k=0}^{n}(f,\varphi_k)\varphi_k,\quad 其中\quad (f,\varphi_k)=\int_{-1}^{1}f(x)\varphi_k(t)\mathrm{d}t.$$

这就是所求的使范数 $\|f-f_n\|$ 最小且次数 $\leqslant n$ 的多项式. 例如, 当 $f(x)=\sin\pi x$ 时, 系数 (f,φ_k) 由

$$(f,\varphi_k)=\int_{-1}^{1}\sin\pi t\varphi_k(t)\mathrm{d}t$$

确定. 特别地, 我们有 $(f,\varphi_0)=0$ 以及

$$(f,\varphi_1)=\int_{-1}^{1}\sqrt{\frac{3}{2}}t\sin\pi t\mathrm{d}t=\sqrt{\frac{3}{2}}\frac{2}{\pi}.$$

因此在区间 $[-1,1]$ 上离 $\sin\pi t$ 最近的线性多项式是

$$f_1(t)=\sqrt{\frac{3}{2}}\frac{2}{\pi}\varphi_1(t)=\frac{3}{\pi}t.$$

由于 $(f,\varphi_2)=0$, 它也是一个最优二次逼近. □

3.17　习　　题

1. 为 $\mathbb{R}^3$ 中由以下各组向量生成的子空间求一组标准正交基.

   (a) $\boldsymbol{x}_1=(1,1,1), \boldsymbol{x}_2=(1,0,1), \boldsymbol{x}_3=(3,2,3)$.

   (b) $\boldsymbol{x}_1=(1,1,1), \boldsymbol{x}_2=(-1,1,-1), \boldsymbol{x}_3=(1,0,1)$.

2. 为 $\mathbb{R}^4$ 中由以下各组向量生成的子空间求一组标准正交基.

   (a) $\boldsymbol{x}_1=(1,1,0,0), \boldsymbol{x}_2=(0,1,1,0), \boldsymbol{x}_3=(0,0,1,1), \boldsymbol{x}_4=(1,0,0,1)$.

   (b) $\boldsymbol{x}_1=(1,1,0,1), \boldsymbol{x}_2=(1,0,2,1), \boldsymbol{x}_3=(1,2,-2,1)$.

   习题 3~10 需要微积分知识.

3. 在实线性空间 $C(0,\pi)$ 中, 设内积为 $(x,y)=\int_0^\pi x(t)y(t)\mathrm{d}t$. 对 $n=0,1,2,\cdots$, 令 $x_n(t)=\cos nt$, 证明由

$$y_0(t)=\frac{1}{\sqrt{\pi}}\quad 和\quad y_n(t)=\sqrt{\frac{2}{\pi}}\cos nt,\quad n\geqslant 1,$$

   确定的函数 $y_0,y_1,y_2,\cdots$ 组成由 $x_0,x_1,x_2,\cdots$ 生成的子空间的一组标准正交组.

4. 在由所有实多项式组成的线性空间中, 设内积为 $(x,y)=\int_0^1 x(t)y(t)\mathrm{d}t$. 对 $n=0,1,2,\cdots$, 令 $x_n(t)=t^n$, 证明函数

$$y_0(t)=1,\quad y_1(t)=\sqrt{3}(2t-1),\quad y_2=\sqrt{5}(6t^2-6t+1),$$

   组成由 $\{x_0,x_1,x_2\}$ 生成的子空间的一组标准正交组.

5. 令 V 为在 $[0,+\infty)$ 上连续且积分 $\int_0^\infty \mathrm{e}^{-t}f^2(t)\mathrm{d}t$ 收敛的所有实函数 f 组成的线性空间, 定义 $(f,g)=\int_0^\infty f(t)g(t)\mathrm{d}t$. 对 $n\geqslant 0$, 令 $x_n(t)=t^n$ 并令 $y_0,y_1,y_2,\cdots$ 为对 $x_0,x_1,x_2,\cdots$ 应用 Gram-Schmidt 方法所得的序列. 证明 $y_0(t)=1$, $y_1(t)=t-1$, $y_2(t)=t^2-4t+2$, $y_3(t)=t^3-9t^2+18t-6$.

6. 在实线性空间 $C(1,3)$ 中, 令内积为 $(f,g)=\int_1^3 f(x)g(x)\mathrm{d}x$. 设 $f(x)=1/x$, 证明离 f 最近的常数函数是 $g=\frac{1}{2}\ln 3$. 计算此时 $\|g-f\|^2$ 的值.

7. 在实线性空间 $C(0,2)$ 中, 令内积为 $(f,g)=\int_0^2 f(x)g(x)\mathrm{d}x$. 设 $f(x)=\mathrm{e}^x$, 证明离 f 最近的常数函数是 $g=\frac{1}{2}(\mathrm{e}^2-1)$. 计算此时 $\|g-f\|^2$ 的值.

8. 在实线性空间 $C(-1,1)$ 中, 令内积为 $(f,g)=\int_{-1}^1 f(x)g(x)\mathrm{d}x$. 设 $f(x)=\mathrm{e}^x$, 求离 f 最近的线性多项式 g, 计算此时 $\|g-f\|^2$ 的值.

9. 在实线性空间 $C(0,2\pi)$ 中, 令内积为 $(f,g)=\int_0^{2\pi} f(x)g(x)\mathrm{d}x$, 设 $f(x)=x$. 在由 $u_0(x)=1$, $u_1(x)=\cos x$, $u_2(x)=\sin x$ 生成的子空间中求离 f 最近的三角多项式.

10. 在习题 5 的线性空间 V 中, 设 $f(x)=\mathrm{e}^{-x}$, 求离 f 最近的线性多项式.

第4章　线性变换·矩阵

4.1　线 性 变 换

系统深入地研究定义域和值域都是线性空间的子集的函数是数学分析的基本目标之一, 我们称这样的函数为**变换** (transformation)、**映射** (mapping) 或**算子** (operator). 本章将讨论最简单的一类函数, 即**线性变换**, 数学的各个分支都要用到这类函数. 通常我们可通过用线性变换逼近的方法获得一般变换的性质.

首先我们介绍关于函数的一些记号和术语. 设 V 和 W 为两个集合, 我们用记号

$$T: V \to W$$

表示 T 是一个定义域为 V 且值域在 W 中的函数. 对 V 中任意 x, 我们称 W 中的元素 $T(x)$ 为 x 在 T 作用下的**象** (image), 并说 T 将 x 映为 $T(x)$. 设 A 为 V 的任意子集, 当 x 取遍 A 中所有元素时, 所得的**象** $T(x)$ 的全体组成的集合称为 A 在 T 作用下的**象**, 并记作 $T(A)$. 定义域 V 的**象** $T(V)$ 称为 T 的**值域** (range).

现在假定 V 和 W 为纯量集相同的两个线性空间, 我们如下定义线性变换.

定义　*设 V 和 W 为两个线性空间, 如果函数 $T: V \to W$ 具有性质:*

(a) $T(x+y) = T(x) + T(y)$, 对 V 中任意 x 和 y;

(b) $T(cx) = cT(x)$, 对 V 中任意 x 以及任意纯量 c,

那么我们称 T 是从 V 到 W 的一个线性变换.

我们可将这两条性质表述为 T 保持加法和纯量乘法运算. 这两条性质可合并成一个公式: 对 V 中任意 x 和 y 以及任意纯量 a 和 b, 都有

$$T(ax+by) = aT(x) + bT(y)$$

由归纳法可得更一般的关系: 对 V 中任意 n 个元素 $x_1, \cdots, x_n$ 以及任意 n 个纯量 $a_1, \cdots, a_n$, 都有

$$T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$$

读者易证以下各例都是线性变换.

例 1 (恒等变换)　若对 V 中任意 x 都有 $T(x) = x$, 则称变换 $T: V \to V$ 为恒等变换, 记作 I 或 I_V. 恒等变换将每个元素都映为它自身. □

例 2 (零变换)　若变换 $T: V \to V$ 将 V 中每个元素都映为 O, 那么 T 称为零变换并记作 O. □

例 3 (与给定纯量 c 的乘积)　即这样的变换 $T:V\to V$, 其中对 V 中任意 x 都有 $T(x)=cx$. 当 $c=1$ 时, 它为恒等变换; 当 $c=0$ 时, 它为零变换. □

例 4 (线性方程组)　设 $V=\mathbb{R}^n$, $W=\mathbb{R}^m$, 给定 mn 个实数 a_{ik}, 其中 $i=1,2,\cdots,m$, $k=1,2,\cdots,n$, 我们定义 $T:\mathbb{R}^n\to\mathbb{R}^m$ 如下: T 通过方程组

$$y_i=\sum_{k=1}^{n}a_{ik}x_k,\quad i=1,2,\cdots,m$$

将 $\mathbb{R}^n$ 中各个向量 $\boldsymbol{x}=(x_1,\cdots,x_n)$ 映为 $\mathbb{R}^m$ 中的向量 $\boldsymbol{y}=(y_1,\cdots,y_m)$. □

例 5 (与给定元素的内积)　设 V 为实欧氏空间, 给定 V 中元素 z, 定义 $T:V\to\mathbb{R}$ 如下: 对任意 $x\in V$, 都有 $T(x)=(x,z)$, 其中 (x,z) 为 x 和 z 的内积. □

例 6 (在子空间上的投影)　设 V 为欧氏空间, S 为 V 的一个有限维子空间且 $e_1,\cdots,e_n$ 为 S 的一组标准正交基, 定义 $T:V\to S$ 如下: 对任意 $x\in V$, 都有 $T(x)=p(x)$, 其中 $p(x)$ 为由

$$px=\sum_{i=1}^{n}(x,e_i)e_i$$

确定的 x 在子空间 S 上的投影. □

下面的两个例子用到微积分的概念.

例 7 (微分算子)　令 V 为在开区间 (a,b) 上可微的所有实函数 f 组成的线性空间, 将 V 中每个函数 f 映为它的导数 f' 的线性变换称为微分算子, 记作 D. 因此我们有 $D:V\to W$, 其中对 V 中任意 f 都有 $D(f)=f'$. 空间 W 包含所有导数 f'. □

例 8 (积分算子)　令 $V=C(a,b)$ 为在区间 $[a,b]$ 上连续的所有实函数组成的线性空间, 对任意 $f\in V$, 令 $g=T(f)$ 为由

$$g(x)=\int_a^x f(t)\mathrm{d}t,\quad a\leqslant x\leqslant b.$$

定义的 V 中函数, 我们称这样的变换 T 为积分算子. □

4.2　零化空间 · 值域

在本节中, 我们用 T 表示线性空间 V 到线性空间 W 的一个线性变换.

定理 4.1　*集合 $T(V)$(T 的值域) 是 W 的子空间. 而且 T 将 V 中的零元素映为 W 中的零元素.*

证明　为了证明 $T(V)$ 是 W 的子空间, 我们只需要验证封闭性公理. 取 $T(V)$ 中任意两个元素, 设为 $T(x)$ 和 $T(y)$, 那么 $T(x)+T(y)=T(x+y)$, 所以 $T(x)+T(y)$ 也是 $T(V)$ 中的元素. 对任意纯量 c, 我们有 $cT(x)=T(cx)$, 所以 $cT(x)$ 也是 $T(V)$ 中的元素, 因此 $T(V)$ 是 W 的子空间. 在关系式 $T(cx)=cT(x)$ 中令 $c=0$ 可知

$T(O) = O$. □

定义 由 V 中被 T 映为 O 的所有元素组成的集合称为 T 的**零化空间** (null space), 记作 $N(T)$. 因此我们有

$$N(T) = \{x : x \in V \quad 且 \quad T(x) = O\}.$$

我们有时也将零化空间称为 T 的**核** (kernel).

定理 4.2 T 的零化空间是 V 的子空间.

证明 如果 x 和 y 在 $N(T)$ 中, 那么由于

$$T(x+y) = T(x) + T(y) = O, \quad T(cx) = cT(x) = O.$$

因此 $x+y$ 和 cx 都在 $N(T)$ 中, 其中 c 为任意纯量. □

以下各例给出了 4.1 节各例中的线性变换的零化空间.

例 1 (恒等变换) 零化空间为 $\{O\}$, 即仅包含 O 的子空间. □

例 2 (零变换) 它把所有元素都映为 O, 所以它的零化空间就是 V. □

例 3 (与给定纯量 c 的乘积) 若 $c \neq 0$, 零化空间为 $\{O\}$; 若 $c = 0$, 零化空间是 V. □

例 4 (线性方程组) 零化空间由 $\mathbb{R}^n$ 中使

$$\sum_{k=1}^{n} a_{ik}x_k = 0, \quad i = 1, 2, \cdots, m.$$

的所有向量 $(x_1, \cdots, x_n)$ 组成. □

例 5 (与给定元素 z 的内积) 零化空间包含 V 中所有与 z 正交的元素. □

例 6 (在子空间 S 上的投影) 对 $x \in V$, 投影 $p(x)$ 等于 $\sum\limits_{i=1}^{n}(x, e_i)e_i$, 其中 $e_1, \cdots, e_n$ 是 S 的一组标准正交基, 因此当且仅当对所有基元素 e_i 都有 $(x, e_i) = 0$ 时, $p(x) = O$, 所以它的零化空间为 $S^{\perp}$, 即 S 的正交补. □

例 7 (微分算子) 零化空间由在给定区间上为常数的所有函数组成. □

例 8 (积分算子) 零化空间仅包含零函数. □

4.3 零化度 · 秩

本节中我们仍然用 T 表示线性空间 V 到线性空间 W 的线性变换. 我们将探索定义域 V、零化空间 $N(T)$ 和值域 $T(V)$ 这几个空间的维数之间的关系.

定义 称零化空间 $N(T)$ 的维数为 T 的**零化度** (nullity). 称值域 $T(V)$ 的维数为 T 的**秩** (rank).

如果 V 是有限维空间, 那么由于零化空间是它的子空间, 所有零化空间也是有限维空间. 下面的定理表明此时值域也是有限维空间.

定理 4.3 (零化度加秩定理)　若 V 的维数有限, 那么 T 的维数也有限, 而且我们有

$$\dim N(T) + \dim T(V) = \dim V. \tag{4.1}$$

也就是说零化度加秩等于定义域的维数.

证明　令 $n = \dim V$, 再令 $e_1, \cdots, e_k$ 为零空间的基, 那么 $k = \dim N(T) \leqslant n$. 由定理 3.7(a), 上述元素是 V 的某组基的一部分, 设这组基为

$$e_1, \cdots, e_k, e_{k+1}, \cdots, e_{k+r}, \tag{4.2}$$

其中 $k + r = n$. 下面我们来证明 r 个元素

$$T(e_{k+1}), \cdots, T(e_{k+r}) \tag{4.3}$$

构成 $T(V)$ 的一组基, 由此即得 $T(V) = r$. 由于 $k + r = n$, 这样我们也证明了公式 (4.1).

首先我们证明 (4.3) 中的 r 个元素生成 $T(V)$. 对任意 $y \in T(V)$, 存在 $x \in V$, 使得 $y = T(x)$, 且存在纯量 $c_1, \cdots, c_{k+r}$ 使得 $x = c_1e_1 + \cdots + c_{k+r}e_{k+r}$, 因此由 $T(e_1) = \cdots = T(e_k) = O$ 可知

$$y = T(x) = \sum_{i=1}^{k+r} c_iT(e_i) = \sum_{i=1}^{k} c_iT(e_i) + \sum_{i=k+1}^{k+r} c_iT(e_i) = \sum_{i=k+1}^{k+r} c_iT(e_i)$$

这证明式 (4.3) 中的元素生成值域 $T(V)$.

下面我们证明这些元素线性无关. 假定可选取纯量 $a_{k+1}, \cdots, a_{k+r}$, 使得

$$\sum_{i=k+1}^{k+r} a_iT(e_i) = O$$

则由此可得

$$T\left(\sum_{i=k+1}^{k+r} a_ie_i\right) = O,$$

所以元素 $z = a_{k+1}e_{k+1} + \cdots + a_{k+r}e_{k+r}$ 在零化空间 $N(T)$ 中. 又因为 $e_1, \cdots, e_k$ 是零化空间的基, 所以存在纯量 $a_1, \cdots, a_k$, 使得 $z = a_1e_1 + \cdots + a_ke_k$. 因此我们有

$$z - z = \sum_{i=1}^{k} a_ie_i - \sum_{i=k+1}^{k+r} a_ie_i = O.$$

又因为式 (4.2) 中的元素线性无关, 所以所有纯量系数 a_i 都必为零, 因此式 (4.3) 中的元素线性无关且组成 $T(V)$ 的一组基.　□

注　若 $N(T)$ 或 $T(V)$ 中至少有一个是无限维空间, 定理 4.3 表明 V 也是一个无限维空间. 反之, 若 V 的维数无限, 那么 $N(T)$ 和 $T(V)$ 的维数中至少有一个无限. 4.4 节中的习题 24 勾画了最后这个命题的证明要点.

4.4 习　题

在习题 1~10 中, 关于 $T(x,y)$ 的公式都定义了一个变换 $T:\mathbb{R}^2\to\mathbb{R}^2$, 其中 (x,y) 为 $\mathbb{R}^2$ 上的点, 判断各题中的 T 是否是线性变换. 如果是, 确定它的零化空间和值域并计算它的零化度和秩.

1. $T(x,y)=(y,x)$.
2. $T(x,y)=(x,-y)$.
3. $T(x,y)=(x,0)$.
4. $T(x,y)=(x,x)$.
5. $T(x,y)=(x^2,y^2)$.
6. $T(x,y)=(\mathrm{e}^x,\mathrm{e}^y)$.
7. $T(x,y)=(x,1)$.
8. $T(x,y)=(x+1,y+1)$.
9. $T(x,y)=(x-y,x+y)$.
10. $T(x,y)=(2x-y,x+y)$.

习题 11 ~ 15 都给出了一个变换 $T:\mathbb{R}^2\to\mathbb{R}^2$, 判断各个 T 是否是线性变换. 如果是, 确定它的零化空间和值域并计算它的零化度和秩.

11. T 将每个点都绕原点旋转 φ 角, 即 T 将极坐标为 (r,θ) 的点映为极坐标为 $(r,\theta+\varphi)$ 的点, 其中 φ 为给定值. 而且 T 将 O 映为 O 本身.
12. T 将每个点都映为该点关于一条过原点的固定直线的反射点.
13. T 将每个点都映为点 $(1,1)$.
14. T 将极坐标为 (r,θ) 的点映到极坐标为 $(2r,\theta)$ 的点, 而且 T 将 O 映为 O 本身.
15. T 将极坐标为 (r,θ) 的点映到极坐标为 $(r,2\theta)$ 的点, 而且 T 将 O 映为 O 本身.

在习题 16~23 中, 关于 $T(x,y,z)$ 的公式都定义了一个变换 $T:\mathbb{R}^3\to\mathbb{R}^3$, 其中 (x,y,z) 为 $\mathbb{R}^3$ 中的点. 判断各个 T 是否是线性变换. 如果是, 确定它的零化空间和值域并计算它的零化度和秩.

16. $T(x,y,z)=(z,y,x)$.
17. $T(x,y,z)=(x,y,0)$.
18. $T(x,y,z)=(x,2y,3z)$.
19. $T(x,y,z)=(x,y,1)$.
20. $T(x,y,z)=(x+1,y+1,z-1)$.
21. $T(x,y,z)=(x+1,y+2,z+3)$.
22. $T(x,y,z)=(x,y^2,z^3)$.
23. $T(x,y,z)=(x+z,0,x+y)$.
24. 令 $T:V\to W$ 为线性空间 V 到线性空间 W 的一个线性变换, 若 V 是无限维的, 证明 $T(V)$ 和 $N(T)$ 中至少有一个是无限维的.

    [提示: 假定 $\dim N(T)=k$, $\dim T(V)=r$, 令 $e_1,\cdots,e_k$ 为 $N(T)$ 的基, $e_1,\cdots,e_k$, $e_{k+1},\cdots,e_{k+n}$ 为 V 中线性无关元素, 其中 $n>r$, 那么由 $n>r$ 可知 $T(e_{k+1}),\cdots,T(e_{k+n})$ 线性相关并由此推出矛盾.]

习题 25~29 都给出了一个变换 $T:V\to V$, 判断各个 T 是否是线性变换. 如果是, 确定它的零化空间和值域并计算它的零化度和秩.

25. 令 V 为次数 $\leqslant n$ 的所有实函数 $p(x)$ 组成的线性空间. 对任意 $p\in V$, $q=T(p)$ 表示对任意实数 x 都有 $q(x)=p(x+1)$.

    习题 26 ~ 30 用到微积分知识.

26. 令 V 为在开区间 $(-1,1)$ 上可微的所有实函数组成的线性空间, 对任意 $f\in V$, $g=T(f)$ 表示对 $(-1,1)$ 中任意 x 都有 $g(x)=xf'(x)$.
27. 令 V 为在开区间 (a,b) 上二阶可微的所有实函数组成的线性空间, 对任意 $y\in V$, 定义 $T(y)=y''+Ay'+By$, 其中 A 和 B 为给定常数.
28. 令 V 为在 $[a,b]$ 上连续的所有实函数组成的线性空间, 对任意 $f\in V$, $g=T(f)$ 表示对 $a\leqslant x\leqslant b$, 都有
$$g(x)=\int_a^b f(t)\sin(x-t)\mathrm{d}t,\quad a\leqslant x\leqslant b.$$
29. 令 V 为所有收敛的实序列 $\{x_n\}$ 组成的线性空间, 定义变换 $T:V\to V$ 如下: 如果 $x=\{x_n\}$ 是极限为 a 的收敛序列, 那么 $T(x)=\{y_n\}$, 其中对 $n\geqslant 1$ 都有 $y_n=a-x_n$.

30. 令 V 为在区间 $[-\pi,\pi]$ 上连续的所有实函数组成的线性空间, 并令 S 为 V 中满足如下三个等式的所有函数 f 组成的 V 的子集:

$$\int_{-\pi}^{\pi} f(t)\mathrm{d}t = 0, \quad \int_{-\pi}^{\pi} f(t)\cos t\mathrm{d}t = 0, \quad \int_{-\pi}^{\pi} f(t)\sin t\mathrm{d}t = 0.$$

(a) 证明 S 是 V 的子空间.

(b) 证明 S 包含所有函数 $f(x) = \cos nx$ 和 $f(x) = \sin nx$, 其中 $n = 2, 3, \cdots$.

(c) 证明 S 是无限维空间.

令 $T: V \to V$ 为如下定义的线性变换: 对任意 $f \in V$, $g = T(f)$ 表示

$$g(x) = \int_{-\pi}^{\pi}\{1 + \cos(x - t)\}f(t)\mathrm{d}t.$$

(d) 证明 T 的值域 $T(V)$ 是有限维空间并求 $T(V)$ 的一组基.

(e) 求 T 的零空间.

(f) 求所有实数 $c \neq 0$ 和 V 中所有非零元素 f 使得 $T(f) = cf$. (注意这样的 f 在 T 的值域中).

4.5 线性变换的代数运算

对值在线性空间 W 中的函数, 可以作如下定义的加法和纯量乘法两种运算.

定义 设 $S: V \to W$ 和 $T: V \to W$ 是两个定义域都为 V 的函数且它们的值都在线性空间 W 中, 若 c 为 w 中的任意纯量, 对 V 中所有 x, 我们分别用等式

$$(S+T)(x) = S(x) + T(x), \quad (cT)(x) = cT(x) \tag{4.4}$$

定义和 $S+T$ 与积 cT.

我们对 V 也是线性空间且它的纯量集与 W 的纯量集相同的情形特别感兴趣. 在这种情形下, 我们用记号 $\mathscr{L}(V,W)$ 表示 V 到 W 的所有线性变换的集合.

如果 S 和 T 是 $\mathscr{L}(V,W)$ 中的两个线性变换, 易证 $S+T$ 和 cT 也是 $\mathscr{L}(V,W)$ 中的线性变换. 更进一步, 在如上定义的运算下, 集合 $\mathscr{L}(V,W)$ 本身就是一个线性空间, 零变换是这个空间中的零元, 变换 $(-1)T$ 是 T 的负元. 易证该空间满足线性空间的所有十条公理, 因此我们有如下定理.

定理 4.4 在由等式 (4.4) 定义的加法和纯量乘法下, 集合 $\mathscr{L}(V,W)$ 是一个线性空间. □

线性变换有一个更为有趣的称为**复合变换** (composition of transformation) 或**变换乘法** (multiplication of transformation) 的代数运算. 这种运算与线性空间的代数结构无关, 所以我们可以将它一般化地定义如下:

定义 (函数的复合) 令 U, V, W 为集合, 设 $T: U \to V$ 是定义域为 U 且值在 V 中的函数; $S: V \to W$ 是定义域为 V 且值在 W 中的函数, 则由等式

$$(ST)(x) = S[T(x)], \quad 任意\ x \in U.$$

定义函数 $ST: U \to W$, 称作 S 与 T 的**复合变换**.

因此为了用 ST 映射 x, 我们首先用 T 映射 x, 再用 S 映射 $T(x)$. 图 4.1 演示

了这个过程.

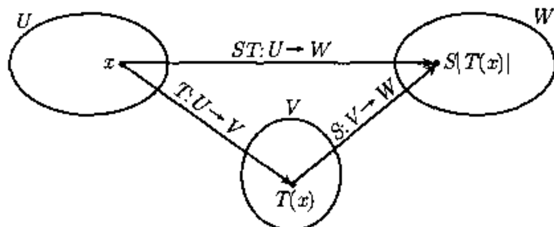


图 4.1 两个变换的复合

在研究初等函数时我们将会碰到实值函数的复合, 此时我们知道, 一般来说, 复合运算不可交换. 例如 $\cos(x^2)$ 表示首先求 x 的平方再求 x^2 的余弦, 它并不总与 $(\cos x)^2$ 相等. $(\cos x)^2$ 表示先求 x 的余弦再求 $\cos x$ 的平方. 尽管如此, 复合运算总是满足结合律.

定理 4.5 如果 $T: U \to V$, $S: V \to W$, $R: W \to X$ 为三个函数, 那么我们有

$$R(ST) = (RS)T.$$

证明 $R(ST)$ 和 $(RS)T$ 的定义域都是 U 且它们的值都在 X 中. 对 U 中任意 x, 我们有

$$[R(ST)](x) = R[(ST)(x)] = R[S\{T(x)\}],$$

$$[(RS)T](x) = (RS)[T(x)] = R[S\{T(x)\}],$$

这证明了 $R(ST) = (RS)T$. □

定义 设 $T: V \to V$ 是将 V 映入它本身的一个函数, 我们归纳地定义 T 的整数幂如下:

$$T^0 = I, \quad T^n = TT^{n-1}, \quad n \geqslant 1.$$

在上述定义中, I 为恒等变换. 由结合律读者易证指数律: 对所有非负整数 m 和 n, 都有 $T^mT^n = T^{m+n}$.

下述定理表明线性变换的复合仍然是线性变换.

定理 4.6 设 U, V, W 为纯量集相同的线性空间, 如果 $T: U \to V$ 和 $S: V \to W$ 都是线性变换, 那么它们的复合 $ST: U \to W$ 也是线性变换.

证明 对 U 中任意 x, y 和任意纯量 a, b, 我们有

$$(ST)(ax + by) = S[T(ax + by)] = S[aT(x) + bT(y)] = aST(x) + bST(y). \quad □$$

我们可将复合运算与 $\mathscr{L}(V, W)$ 中的加法和纯量乘法运算结合在一起得到如下定理.

定理 4.7 令 U, V, W 为纯量集相同的线性空间, 设 S 和 T 属于 $\mathscr{L}(V, W)$, 并令 c 为任意纯量, 则

(a) 对值在 V 中的任意函数 R, 我们有

$$(S+T)R = SR + TR, \quad (cS)R = c(SR).$$

(b) 对任意线性变换 $R: W \to U$, 我们有

$$R(S+T) = RS + RT, \quad R(cS) = c(RS).$$ □

由复合运算的定义易证本定理, 它的证明留给读者作为练习.

4.6 逆

在对初等函数的研究中, 我们证明了可通过对单调函数求反函数来构造新函数. 例如, 自然对数函数 $f(x) = \ln x$ 对 $x > 0$ 单调增, 它的反函数是指数函数, 即对 $x > 0$, 当且仅当 $x = \mathrm{e}^y$ 时, $y = \ln x$. 本节将反函数概念推广到更广泛的一类函数中.

给定函数 T (可以不是线性的), 我们的目的是在可能的情况下找到另一个函数 S, 使得它与 T 的复合为恒等变换. 由于在一般意义下, 复合运算不可交换, 我们需要区别 ST 和 TS, 因此我们引入两种逆, 分别称这两种逆为左逆和右逆.

定义 给定两个集合 V 和 W 以及函数 $T: V \to W$, 函数 $S: T(V) \to V$, $R: T(V) \to V$, 如果对 V 中任意 x, 都有 $S[T(x)] = x$, 即如果

$$ST = I_V,$$

其中 I_V 为 V 上的恒等变换, 那么我们称 S 为 T 的**左逆** (left inverse).

如果对 $T(V)$ 中任意 y, $T[R(y)] = y$, 即如果

$$TR = I_{T(V)},$$

其中 $I_{T(V)}$ 为 $T(V)$ 上的恒等变换, 那么我们称 R 为 T 的**右逆**(right inverse).

例 (没有左逆但有两个右逆的函数) 令 $V = \{1, 2\}$, $W = \{0\}$, 定义 $T: V \to W$ 如下: $T(1) = T(2) = 0$, 这个函数有如下两个右逆 R 和 R':

$$R(0) = 1, \quad R'(0) = 2.$$

如果它有左逆, 那么要求

$$1 = S[T(1)] = S(0), \quad 2 = S[T(2)] = S(0),$$

这与函数值 $S(0)$ 必须唯一矛盾, 所以它不可能有左逆. 这个简单的例子表明左逆不一定存在, 右逆不一定唯一. □

易证任意函数 $T: U \to W$ 都至少有一个右逆. 事实上, 对 $T(V)$ 中任意 y, 都至少存在 V 中的一个 x 使得 $y = T(x)$, 如果我们选定一个这样的 x 并定义 $R(y) = x$, 则对 $T(V)$ 中任意 y, 都有 $T[R(y)] = T(x) = y$, 所以 R 是 T 的右逆. 由于可能有多个 V 中的 x 被映射到 $T(V)$ 中一个给定的 y, 所以右逆可能不唯一. 我们马上就会证明 (定理 4.9), 如果 $T(V)$ 中的每个 y 都恰好是 V 中唯一一个 x 的象, 那么右逆唯一.

首先我们证明: 如果左逆存在, 那么它唯一且它也是右逆.

定理 4.8 函数 $T:V\to W$ 最多有一个左逆. 如果 T 有左逆 S, 那么 S 也是 T 的右逆.

证明 假定 T 有两个左逆 $S:T(V)\to V$ 和 $S':T(V)\to V$, 取 $T(V)$ 中任意 y, 我们需要证明 $S(y)=S'(y)$. 由于存在 V 中 x 使得 $y=T(x)$ 且 S 和 S' 都是 T 的左逆, 所以我们有

$$S[T(x)]=x, \quad S'[T(x)]=x.$$

这表示 $S(y)=x$ 且 $S'(y)=x$, 所以对 $T(V)$ 中任意 y, 都有 $S(y)=S'(y)$, 因此 $S=S'$, 所以如果左逆存在那么它唯一.

现在我们来证明任意左逆 S 也都是右逆. 任取 $T(V)$ 中 y, 我们需要证明 $T[S(y)]=y$. 由于 $y\in T(V)$, 所以存在 V 中 x, 使得 $y=T(x)$. 又因为 S 是左逆, 所以

$$x=S[T(x)]=S(y).$$

将 T 同时作用于上式两边, 我们有 $T(x)=T[S(y)]$. 又因为 $T(x)=y$, 所以 $y=T[S(y)]$, 因此定理得证. □

这样一来描述有左逆的所有函数就变得十分简单. 首先我们引入如下术语.

定义 当函数 $T:V\to W$ 将 V 中不同元素映射成 W 中不同元素时, 我们称 T 在 V 上一一 (one-to-one), 即如果对 V 中任意 x,y, 都有

$$x\neq y \quad \text{必然包含} \quad T(x)\neq T(y). \tag{4.5}$$

我们称 T 在 V 上一一.

下述命题与 (4.5) 等价:

$$T(x)=T(y) \quad \text{必然包含} \quad x=y. \tag{4.6}$$

定理 4.9 当且仅当函数 $T:V\to W$ 在 V 上一一时, 它有左逆.

证明 假定 T 有左逆 S 且存在 S 中两个元素 x 和 y, 使得 $T(x)=T(y)$, 将 S 同时作用于此式两边, 我们有 $x=y$, 所以 T 在 V 上一一.

对逆命题, 假定 T 在 V 上一一, 我们需要构造一个函数 $S:T(V)\to V$ 且它是 T 的左逆. 对任意 $y\in T(V)$, 由于 T 在 V 上一一, 所以有且仅有一个 V 中的 x 使得 $y=T(x)$. 将这个 x 定义为 $S(y)$ 的值, 即我们在 $T(V)$ 上定义 S 如下:

$$S(y)=x \quad \text{表示} \quad T(x)=y.$$

那么对 V 中任意 x, 都有 $S[T(x)]=S(y)=x$, 所以 S 是 T 的左逆. □

定义 (可逆函数) 设 $T:V\to W$ 在 V 上一一, 我们将 T 的唯一左逆 (也是 T 的唯一右逆) 记作 T^{-1}, 我们说 T 可逆, 且称 T^{-1} 为 T 的逆.

本节的结果对任意函数都成立. 下一节我们将会把这些结果应用到线性变换中.

4.7　一一线性变换

本节中, V 和 W 为纯量集相同的两个线性空间, $T: V \to W$ 为 $\mathscr{L}(V,W)$ 中的线性变换. T 的线性性使我们可将它的一一性转化成多种等价形式.

定理 4.10　设 $T: V \to W$ 为 $\mathscr{L}(V,W)$ 中的线性变换, 那么下述各命题等价.

(a) T 在 V 上一一.

(b) T 可逆且它的逆 $T^{-1}: T(V) \to V$ 也是线性变换.

(c) $\dim N(T) = 0$, 即 T 的零空间 $N(T)$ 仅包含零元素.

证明　我们将会证明由 (a) 可得 (b), 由 (b) 可得 (c), 由 (c) 可得 (a). 首先假定 (a) 成立, 那么 T^{-1} 存在 (由定理 4.9), 我们需要证明它是线性变换. 任取 $T(V)$ 中两元素 u 和 v, 则存在 V 中元素 x 和 y 使得 $u = T(x)$, $v = T(y)$. 因为 T 是线性变换, 所以对任意纯量 a 和 b, 有

$$au + bv = aT(x) + bT(y) = T(ax + by),$$

将 T^{-1} 同时作用于上式两边可得

$$T^{-1}(au + bv) = ax + by = aT^{-1}(u) + bT^{-1}(v),$$

所以 T^{-1} 是线性变换, 因此由 (a) 可得 (b).

接下来假定 (b) 成立, 任取 V 中使 $T(x) = O$ 的元素 x, 将 T^{-1} 同时作用于此式两边且由于 T^{-1} 是线性变换, 所以我们有 $x = T^{-1}(O) = O$, 因此由 (b) 可得 (c).

随后假定 (c) 成立, 任取 V 中两个函数值相等的元素 x 和 y, 则 $T(x) = T(y)$. 由 T 的线性性可知 $T(x - y) = T(x) - T(y) = O$, 所以 $x - y = O$, 即 $x = y$, 因此 T 在 V 上一一, 所以由 (c) 可得 (a). 定理得证. □

当 V 为有限维空间时, 我们可用线性无关性和维数的公式表示一一性, 下述定理给出了这种表示法.

定理 4.11　设 $T: V \to W$ 为 $\mathscr{L}(V,W)$ 中的线性变换, 假定 V 是有限维空间, 设 $\dim V = n$, 那么下列各命题等价.

(a) T 在 V 上一一.

(b) 若 $v_1, \cdots, v_p$ 为 V 中 p 个线性无关元素, 那么 $T(v_1), \cdots, T(v_p)$ 为 $T(V)$ 中的 p 个线性无关元素.

(c) $\dim T(V) = \dim V$.

(d) 若 $\{v_1, \cdots, v_n\}$ 是 V 的基, 那么 $\{T(v_1), \cdots, T(v_n)\}$ 是 $T(V)$ 的基.

证明　我们将会证明由 (a) 可得 (b), 由 (b) 可得 (c), 由 (c) 可得 (d), 由 (d) 可得 (a). 首先假定 (a) 成立, 令 $v_1, \cdots, v_p$ 为 V 中线性无关元素, 考虑 $T(V)$ 中元素 $T(v_1), \cdots, T(v_p)$. 设纯量 $c_1, \cdots, c_p$ 使 $\sum_{i=1}^{p} c_i T(v_i) = O$, 因为 T 是线性变换且在 V 上一一, 所以我们有

$$T\left(\sum_{i=1}^{p} c_i v_i\right) = O, \quad 因此 \quad \sum_{i=1}^{p} c_i v_i = O.$$

又因为 $v_1, \cdots, v_p$ 线性无关, 所以 $c_1 = \cdots = c_p = 0$, 因此由 (a) 可得 (b).

现在假设 (b) 成立, 如果 $\{v_1, \cdots, v_n\}$ 是 V 的基, 由 (b) 可知 $T(V)$ 中的 n 个元素 $T(v_1), \cdots, T(v_n)$ 线性无关, 因此 $\dim T(V) \geqslant n$. 另一方面, 由定理 4.3 可知 $\dim T(V) \leqslant n$, 所以 $\dim T(V) = n$, 因此由 (b) 可得 (c).

接下来假定 (c) 成立, 令 $\{v_1, \cdots, v_n\}$ 为 V 的基, 任取 $T(V)$ 中元素 y, 那么存在 V 中元素 x 使 $y = T(x)$, 由此我们可将 x 写成

$$x = \sum_{i=1}^{n} c_i v_i, \quad 因此 \quad y = T(x) = \sum_{i=1}^{n} c_i T(v_i).$$

于是元素集 $\{T(v_1), \cdots, T(v_n)\}$ 生成 $T(V)$. 又因为 $\dim T(V) = n$, 所以 $\{T(v_1), \cdots, T(v_n)\}$ 是 $T(V)$ 的基, 因此由 (c) 可得 (d).

最后假定 (d) 成立, 我们将会证明由 $T(x) = O$ 可得 $x = O$. 令 $\{v_1, \cdots, v_n\}$ 为 V 的基, 对任意 $x \in V$, 我们可将它写作

$$x = \sum_{i=1}^{n} c_i v_i, \quad 因此 \quad T(x) = \sum_{i=1}^{n} c_i T(v_i).$$

如果 $T(x) = O$, 因为元素 $T(v_1), \cdots, T(v_n)$ 线性无关, 所以 $c_1 = \cdots = c_n = 0$. 于是 $x = 0$, 因此 T 在 V 上一一, 从而由 (d) 可得 (a). 定理得证. □

4.8 习　　题

1. 令 $V = \{0, 1\}$, 确定所有函数 $T : V \to V$. 一共有四个这样的函数, 将它们记作 T_1, T_2, T_3, T_4, 做一个乘法表表示它们的两两复合, 指出哪些函数在 V 上一一并给出它们的逆.
2. 令 $V = \{0, 1, 2\}$, 确定所有使得 $T(V) = V$ 的函数 $T : V \to V$. 一共有六个这样的函数, 将它们记作 $T_1, \cdots, T_6$, 做一个乘法表表示它们的两两复合, 指出哪些函数在 V 上一一并给出它们的逆.

在习题 3~12 中, 函数 $T : \mathbb{R}^2 \to \mathbb{R}^2$ 都由一个关于 $T(x, y)$ 的公式所定义, 其中 (x, y) 为 $\mathbb{R}^2$ 中的任意点, 判断各函数 T 是否在 $\mathbb{R}^2$ 上一一, 如果是, 设 $(x, y) = T^{-1}(u, v)$, 给出用 u 和 v 表示 x 和 y 的公式.

3. $T(x, y) = (y, x)$.
4. $T(x, y) = (x, -y)$.
5. $T(x, y) = (x, 0)$.
6. $T(x, y) = (x, x)$.
7. $T(x, y) = (x^2, y^2)$.
8. $T(x, y) = (\mathrm{e}^x, \mathrm{e}^y)$.
9. $T(x, y) = (x, 1)$.
10. $T(x, y) = (x + 1, y + 1)$.
11. $T(x, y) = (x - y, x + y)$.
12. $T(x, y) = (2x - y, x + y)$.

在习题 13~20 中, 函数 $T : \mathbb{R}^3 \to \mathbb{R}^3$ 都由一个关于 $T(x, y, z)$ 的公式所定义, 其中 (x, y, z) 为 $\mathbb{R}^3$ 中的任意点, 判断各函数 T 是否在 $\mathbb{R}^3$ 上一一, 如果是, 设 $(x, y, z) = T^{-1}(u, v, w)$, 给出用 u, v, w 表示 x, y, z 的公式.

13. $T(x, y, z) = (z, y, x)$.
14. $T(x, y, z) = (x, y, 0)$.

15. $T(x, y, z) = (x, 2y, 3z)$.
16. $T(x, y, z) = (x, y, x + y + z)$.
17. $T(x, y, z) = (x + 1, y + 1, z - 1)$.
18. $T(x, y, z) = (x + 1, y + 2, z + 3)$.
19. $T(x, y, z) = (x, x + y, x + y + z)$.
20. $T(x, y, z) = (x + y, y + z, x + z)$.
21. 令 $T : V \to V$ 为将 V 映到它自身的一个函数, 用公式 $T^0 = I$, 以及对 $n \geqslant 1$, $T^n = TT^{n-1}$ 归纳地定义 T 的幂, 证明由复合的结合律可得指数定律: $T^m T^n = T^{m+n}$. 如果 T 可逆, 证明 T^n 也可逆且 $(T^n)^{-1} = (T^{-1})^n$.

在习题 22~25 中, S 和 T 表示定义域为 V 且值在 V 中的函数, 一般来说 $ST \neq TS$, 如果 $ST = TS$, 那么我们称 S 和 T **可交换** (commute).

22. 若 S 和 T 可交换, 证明对任意整数 $n \geqslant 0$, 都有 $(ST)^n = S^n T^n$.
23. 若 S 和 T 可逆, 证明 ST 也可逆且 $(ST)^{-1} = T^{-1}S^{-1}$, 也就是说, 复合的逆等于逆的逆序复合.
24. 若 S 和 T 可逆且可交换, 证明它们的逆也可交换.
25. 设 V 为线性空间, 若 S 和 T 可交换, 证明

$$(S + T)^2 = S^2 + 2ST + T^2, \quad (S + T)^3 = S^3 + 3S^2T + 3ST^2 + T^3.$$

当 $ST \neq TS$ 时, 指出应该对这两个公式做什么变动从而使它们成立.

26. 设 S 和 T 属于 $\mathscr{L}(V, V)$, 假定 $ST - TS = I$, 证明对任意 $n \geqslant 1$, 都有 $ST^n - T^nS = nT^{n-1}$.
27. 设 S 和 T 分别是由公式 $S(x, y, z) = (z, y, x)$ 和 $T(x, y, z) = (x, x + y, x + y + z)$ 定义的从 $\mathbb{R}^3$ 到 $\mathbb{R}^3$ 的线性变换, 其中 (x, y, z) 为 $\mathbb{R}^3$ 中的任意点,
   (a) 求在下述各变换下 (x, y, z) 的象: $ST, TS, ST - TS, S^2, T^2, (ST)^2, (TS)^2, (ST - TS)^2$.
   (b) 证明 S 和 T 在 $\mathbb{R}^3$ 上一一并求在下述各变换下 (u, v, w) 的象: $S^{-1}, T^{-1}, (ST)^{-1}, (TS)^{-1}$.
   (c) 对每个 $n \geqslant 1$, 求 (x, y, z) 在 $(T - I)^n$ 下的象.
28. 设 V 为所有实多项式 $p(x)$ 组成的线性空间, 令 R, S, T 分别是将多项式 $p(x) = c_0 + c_1x + \cdots + c_nx^n$ 映射为多项式 $r(x), s(x)$ 和 $t(x)$ 的函数, 其中

$$r(x) = p(0), \quad s(x) = \sum_{k=1}^{n} c_k x^{k-1}, \quad t(x) = \sum_{k=0}^{n} c_k x^{k+1}.$$

   (a) 令 $p(x) = 2 + 3x - x^2 + x^3$, 求在下述各变换下 p 的象: $R, S, T, ST, TS, (TS)^2, T^2S^2, S^2T^2, TRS, RST$.
   (b) 证明 R, S, T 都是线性变换并分别求它们的零空间和值域.
   (c) 证明 T 在 V 上一一并求它的逆.
   (d) 若 $n \geqslant 1$, 用 I 和 R 表示 $(TS)^n$ 和 S^nT^n.

习题 29~32 需要微积分知识.

29. 设 V 为所有实函数 $p(x)$ 组成的线性空间, 令 D 为微分算子, 令 T 表示将 $p(x)$ 映为 $xp'(x)$ 的线性变换,
   (a) 令 $p(x) = 2 + 3x - x^2 + 4x^3$, 求在下述各变换下 p 的象: $D, T, DT, TD, DT - TD, T^2D^2 - D^2T^2$.
   (b) 求 V 中使 $T(p) = p$ 的所有 p.
   (c) 求 V 中使 $(DT - 2D)(p) = O$ 的所有 p.
   (d) 求 V 中使 $(DT - TD)^n(p) = D^n(p)$ 的所有 p.
30. 设 V 和 D 是如习题 29 所述的函数, 令 T 为将 $p(x)$ 映射为 $xp(x)$ 的线性变换, 证明 $DT - TD = I$ 且对 $n \geqslant 2$, 都有 $DT^n - T^nD = nT^{n-1}$.
31. 令 V 为所有实多项式 $p(x)$ 组成的线性空间, 令 D 表示微分算子并令 T 表示积分算子, 即 T 将每一个多项式 p 映为多项式 q, 其中 $q(x) = \int_0^x p(t)\mathrm{d}t$, 证明 $DT = I_V$ 但 $TD \neq I_V$. 求 TD 的零空间和值域.
32. 参见 4.4 节中习题 29, 判断 T 是否在 V 上一一, 如果是, 求它的逆.

4.9 基元素的象为指定值的线性变换

下一条定理表明，如果 V 是有限维空间，那么线性变换 $T:V\to V$ 由它在 V 的基元素上的作用完全确定.

定理 4.12 设 $v_1,\cdots,v_n$ 为 n 维线性空间 V 的一组基，令 $u_1,\cdots,u_n$ 为线性空间 W 中的 n 个任意元素，那么有且仅有一个线性变换 $T:V\to W$ 使得对 $k=1,2,\cdots,n$，都有

$$T(v_k)=u_k. \tag{4.7}$$

对 V 中任意元素 x，若 $x=\sum_{k=1}^{n}x_kv_k$，则 T 将 x 映为

$$T(x)=\sum_{k=1}^{n}x_ku_k. \tag{4.8}$$

证明 V 中任意元素 x 都是 $v_1,\cdots,v_n$ 的线性组合，该线性组合中的系数 $x_1,\cdots,x_n$ 是 x 关于有序基 $(v_1,\cdots,v_n)$ 的分量. 如果我们按式 (4.8) 定义 T，易证 T 是线性变换. 若 x 等于基中某一元素 v_k，那么 x 的第 k 个分量为 1，其余分量都为 0，所以由式 (4.8) 得 $T(v_k)=u_k$.

为了证明仅有一个线性变换满足条件 (4.7)，令 T' 为另一个这样的变换，计算 $T'(x)$，我们可得

$$T'(x)=T'\left(\sum_{k=1}^{n}x_kv_k\right)=\sum_{k=1}^{n}x_kT'(v_k)=\sum_{k=1}^{n}x_ku_k=T(x).$$

由于对 V 中任意 x，都有 $T'(x)=T(x)$，所以我们有 $T'=T$. 定理得证. □

例 求将基元素 $\boldsymbol{i}=(1,0)$ 和 $\boldsymbol{j}=(0,1)$ 映为

$$T(\boldsymbol{i})=\boldsymbol{i}+\boldsymbol{j},\quad T(\boldsymbol{j})=2\boldsymbol{i}-\boldsymbol{j}.$$

的线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$.

解 设 $\boldsymbol{x}=x_1\boldsymbol{i}+x_2\boldsymbol{j}$ 为 $\mathbb{R}^2$ 中任意元素，则 $T(x)$ 由

$$T(\boldsymbol{x})=x_1T(\boldsymbol{i})+x_2T(\boldsymbol{j})=x_1(\boldsymbol{i}+\boldsymbol{j})+x_2(2\boldsymbol{i}-\boldsymbol{j})=(x_1+2x_2)\boldsymbol{i}+(x_1-x_2)\boldsymbol{j}.$$

确定. □

4.10 线性变换的矩阵表示

定理 4.12 表明有限维线性空间 V 上的线性变换 $T:V\to W$ 完全由它在给定的一组基 $v_1,\cdots,v_n$ 上的作用确定. 现在假定 W 也是有限维空间，设 $\dim W=m$，令 $w_1,\cdots,w_m$ 为 W 的基 (维数 n 和 m 可能相等也可能不等). 由于 T 的值在 W 中，所以任意元素 $T(v_k)$ 都可唯一地表示为 $w_1,\cdots,w_m$ 的线性组合，设

$$T(v_k)=\sum_{i=1}^{m} t_{ik}w_i,$$

其中 $t_{1k},\cdots,t_{mk}$ 为 $T(v_k)$ 关于有序基 $(w_1,\cdots,w_m)$ 的分量, 我们将 m 元组 $(t_{1k},\cdots,t_{mk})$ 写成如下竖直形式:

$$\begin{bmatrix} t_{1k} \\ t_{2k} \\ \vdots \\ t_{mk} \end{bmatrix} \tag{4.9}$$

称这样的数组为**列向量**或**列矩阵**. 对 $T(v_1),\cdots,T(v_n)$ 中的每一个元素, 我们都有一个这样的列向量. 若我们将这些列向量并列排好并用一对中括号把它们括起来, 那么我们将得到如下矩形阵列 (array):

$$\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix}$$

我们称这样的阵列为由 m **行** n **列组成的矩阵**, 并记作 $m\times n$ 矩阵 (读作 m 行 n 列矩阵). 它的第一行是一个 $1\times n$ 矩阵 $(t_{11},t_{12},\cdots,t_{1n})$, 而式 (4.9) 中所示的 $m\times 1$ 矩阵是它的第 k 列. 纯量 t_{ik} 的下标表示它在矩阵中的位置, 其中第一个下标 i 表示它所在的行, 第二个下标 k 表示它所在的列. 我们称 t_{ik} 为该矩阵中的第 ik 个**元** (entry) 或第 ik 个**元素** (element). 也可用更紧凑的记号

$$(t_{ik}) \quad 或 \quad (t_{ik})_{i,k=1}^{m,n}$$

表示同样的矩阵, 这种形式表示 t_{ik} 是矩阵的第 ik 个元, 也用 $[t_{ik}]$ 表示这个矩阵.

因此 n 维空间 V 到 m 维空间 W 上的任意线性变换都确定一个 $m\times n$ 矩阵 (t_{ik}), 该矩阵的各列包含 $T(v_1),\cdots,T(v_n)$ 关于基 $(w_1,\cdots,w_m)$ 的各分量. 我们称该矩阵为 T 关于 V 中给定的有序基 $(v_1,\cdots,v_n)$ 和 W 中有序基 $(w_1,\cdots,w_m)$ 的**矩阵表示** (matrix representation) 或**矩阵**. 一旦我们知道矩阵 (t_{ik}), 按下述定理可确定任意元素 $T(x)$ 关于基 $(w_1,\cdots,w_m)$ 的分量:

定理 4.13 设 T 为 $\mathscr{L}(V,W)$ 中的线性变换, 其中 $\dim V=n$, $\dim W=m$. 令 $(v_1,\cdots,v_n)$ 和 $(w_1,\cdots,w_m)$ 分别为 V 和 W 的有序基, 并令 $m\times n$ 矩阵 (t_{ik}) 的元由等式

$$T(v_k)=\sum_{i=1}^{m} t_{ik}w_i, \quad k=1,2,\cdots,n. \tag{4.10}$$

确定, 那么 T 将 V 中关于基 $(v_1,\cdots,v_n)$ 的分量为 $(x_1,\cdots,x_n)$ 的任意元素

$$x = \sum_{k=1}^{n} x_k v_k \tag{4.11}$$

映为 W 中关于基 $(w_1, \cdots, w_m)$ 的分量为 $(y_1, \cdots, y_m)$ 的元素

$$T(x) = \sum_{i=1}^{m} y_i w_i \tag{4.12}$$

其中 y_i 关于 x 的分量的关系由线性方程

$$y_i = \sum_{k=1}^{n} t_{ik} x_k, \quad i = 1, 2, \cdots, m. \tag{4.13}$$

确定.

证明 把 T 同时作用于等式 (4.11) 两边并由关系式 (4.10) 可得

$$T(x) = \sum_{k=1}^{n} x_k T(v_k) = \sum_{k=1}^{n} x_k \sum_{i=1}^{m} t_{ik} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} t_{ik} x_k \right) w_i = \sum_{i=1}^{m} y_i w_i,$$

其中所有 y_i 都由式 (4.13) 确定. 定理得证. □

一旦我们分别选定 V 和 W 的基 $(v_1, \cdots, v_n)$ 和 $(w_1, \cdots, w_m)$, 任意线性变换 $T: V \to W$ 都有一个完全确定的矩阵表示 (t_{ik}). 反之, 如果我们先选定 mn 个纯量并将它们排列成一个矩阵 (t_{ik}), 再分别为 V 和 W 选定一组有序基, 那么易证恰有一个以 (t_{ik}) 为矩阵表示的线性变换 $T: V \to W$. 我们按 (4.10) 中的等式定义 T 在 V 的基元素上的值, 那么由定理 4.12, 有且仅有一个使基元素的象分别为这些值的线性变换 $T: V \to W$. V 中任意元素 x 的象 $T(x)$ 由等式 (4.12) 和 (4.13) 确定.

例 1 (由给定矩阵构造线性变换) 假定我们有 2×3 矩阵

$$\begin{bmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \end{bmatrix}$$

选 $\mathbb{R}^3$ 和 $\mathbb{R}^2$ 的单位坐标向量分别作为它们的基, 那么给定矩阵表示线性变换 $T: \mathbb{R}^3 \to \mathbb{R}^2$, T 按线性方程组

$$y_1 = 3x_1 + x_2 - 2x_3,$$
$$y_2 = x_1 + 0x_2 - 4x_3.$$

将 $\mathbb{R}^3$ 中向量 (x_1, x_2, x_3) 映为 $\mathbb{R}^2$ 中向量 (y_1, y_2). □

例 2 (求给定线性变换的矩阵表示) 令 $V = W = \mathbb{R}^2$ 且取单位坐标向量作为它们的基, 令 T 为按线性方程组

$$y_1 = 2x_1 + 3x_2,$$
$$y_2 = 4x_1 + 5x_2.$$

将任意向量 (x_1, x_2) 映为向量 (y_1, y_2) 的线性变换, 可用 2×2 矩阵

$$\begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix}$$

表示该变换. □

例 3 取 $(\boldsymbol{i},\boldsymbol{j})$ 为 $\mathbb{R}^2$ 的基, 设线性变换 $T:\mathbb{R}^2\to\mathbb{R}^2$ 将 $(1,1)$ 映射为 $(1,\frac{1}{2})$, 将 $(1,-1)$ 映射为 $(\frac{1}{2},1)$, 求 T 的矩阵表示.

解 由给定条件可确定 T 在基元素 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 上的作用. 给定条件告诉我们

$$T(\boldsymbol{i}+\boldsymbol{j})=\boldsymbol{i}+\frac{1}{2}\boldsymbol{j},\quad T(\boldsymbol{i}-\boldsymbol{j})=\frac{1}{2}\boldsymbol{i}+\boldsymbol{j}.$$

由 T 的线性性我们有 $T(\boldsymbol{i}+\boldsymbol{j})=T(\boldsymbol{i})+T(\boldsymbol{j})$ 和 $T(\boldsymbol{i}-\boldsymbol{j})=T(\boldsymbol{i})-T(\boldsymbol{j})$, 所以由给定条件可得

$$T(\boldsymbol{i})+T(\boldsymbol{j})=\boldsymbol{i}+\frac{1}{2}\boldsymbol{j},$$
$$T(\boldsymbol{i})-T(\boldsymbol{j})=\frac{1}{2}\boldsymbol{i}+\boldsymbol{j}.$$

将以上两式相加可得

$$T(\boldsymbol{i})=\frac{3}{4}\boldsymbol{i}+\frac{3}{4}\boldsymbol{j}$$

相减可得

$$T(\boldsymbol{j})=\frac{1}{4}\boldsymbol{i}-\frac{1}{4}\boldsymbol{j}$$

取两式中 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 的系数作为列, 我们可知 T 的矩阵为

$$\begin{bmatrix}\frac{3}{4} & \frac{1}{4}\\ \frac{3}{4} & -\frac{1}{4}\end{bmatrix}$$

与 T 对应的线性方程组为

$$y_1=\frac{3}{4}x_1+\frac{1}{4}x_2,\quad y_2=\frac{3}{4}x_1-\frac{1}{4}x_2.$$ □

上述各例表明了线性变换的矩阵表示与建立各分量间关系的线性方程组之间的关系. 在以下两个例子中, 我们将用几何方法取代线性方程来表示线性变换. 在这两个例子中, 我们取 $V=W=\mathbb{R}^n$, 其中 $n=2$(例 4) 或 3(例 5), 并取单位坐标向量作为 V 和 W 的基. 和以前一样, 我们通过观察 T 在基向量上的作用从而确定它的矩阵表示. 由矩阵表示易得表示变换前后各分量间关系的线性方程.

例 4 (平面上的旋转) 令 $T:\mathbb{R}^2\to\mathbb{R}^2$ 表示关于一条过原点且垂直于平面 $\mathbb{R}^2$ 的轴的旋转, 它的旋转角为 θ. 为了确定 T 的矩阵表示, 我们考虑 T 在基元素 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 上的作用. 旋转 θ 角意味着将点 $\boldsymbol{i}=(1,0)$ 映射为 $T(\boldsymbol{i})=(\cos\theta,\sin\theta)$ 且将点 $\boldsymbol{j}=(0,1)$ 映射为 $(\cos(\theta+\frac{1}{2}\pi),\sin(\theta+\frac{1}{2}\pi))=(-\sin\theta,\cos\theta)$, 用基元素表示这两个象, 我们有

$$T(\boldsymbol{i})=\cos\theta\ \boldsymbol{i}+\sin\theta\ \boldsymbol{j}$$
$$T(\boldsymbol{j})=-\sin\theta\ \boldsymbol{i}+\cos\theta\ \boldsymbol{j}.$$

由前述例题, 我们知道象 $T(\boldsymbol{i})$ 和 $T(\boldsymbol{j})$ 的分量分别构成 T 的矩阵表示的各列, 所以 T 由 2×2 矩阵

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

表示, 我们称该矩阵为**旋转矩阵**(rotation matrix). 有了 T 的矩阵表示, 我们易求描述该旋转的线性方程. 旋转角 θ 将点 (x_1, x_2) 映射为 (y_1, y_2), 其中

$$\begin{aligned} y_1 &= x_1\cos\theta - x_2\sin\theta, \\ y_2 &= x_1\sin\theta + x_2\cos\theta. \end{aligned}$$ □

例 5 (三维空间中的旋转) 令 $T: \mathbb{R}^3 \to \mathbb{R}^3$ 表示关于 x_1 轴的旋转且它的旋转角为 θ, x_1 轴上的点不变, 然而 x_2x_3 平面中的点将旋转 θ 角, 对 x_2x_3 平面应用例 4, 我们知道基元素 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 将进行如下映射:

$$\begin{aligned} T(\boldsymbol{i}) &= \boldsymbol{i} = 1\boldsymbol{i} + 0\boldsymbol{j} + 0\boldsymbol{k} \\ T(\boldsymbol{j}) &= 0\boldsymbol{i} + \cos\theta\boldsymbol{j} + \sin\theta\boldsymbol{k} \\ T(\boldsymbol{k}) &= 0\boldsymbol{i} - \sin\theta\boldsymbol{j} + \cos\theta\boldsymbol{k}. \end{aligned}$$

这些分量组成 T 的如下 3×3 矩阵表示的各列:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

T 将点 (x_1, x_2, x_3) 映射为点 (y_1, y_2, y_3), 其中

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= x_2\cos\theta - x_3\sin\theta \\ y_3 &= x_2\sin\theta + x_3\cos\theta. \end{aligned}$$ □

接下来的例子需要微积分知识

例 6 令 V 为次数 $\leqslant 3$ 的所有实多项式 $p(x)$ 组成的线性空间, 该空间的维数为 4, 选定它的有序基 $(1, x, x^2, x^3)$. 令 D 为微分算子, 它将 V 中任意多项式 $p(x)$ 映为它的导数 $p'(x)$. 我们可将 D 看作从 V 到 W 的一个线性变换, 其中 W 为次数 $\leqslant 2$ 的所有实多项式组成的三维空间. 在 W 中取有序基 $(1, x, x^2)$, 为了求 D 关于选定的基的矩阵表示, 我们对 V 的每一个基元素作变换 (求导) 并将所得结果表示成 W 的基元素的线性组合. 因此我们有

$$D(1) = 0 = 0 + 0x + 0x^2, \quad D(x) = 1 = 1 + 0x + 0x^2,$$
$$D(x^2) = 2x = 0 + 2x + 0x^2, \quad D(x^3) = 3x^2 = 0 + 0x + 3x^2.$$

这些多项式中的系数确定了 D 的矩阵表示的各列, 所以如下 3×4 矩阵是 D 的矩阵表示:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

为了表明变换的矩阵表示不仅依赖于所选的基元素, 还依赖于基元素的顺序, 我们将 W 的基元素逆序排列, 即取有序基 $(x^2, x, 1)$, 那么 V 的基元素仍被变换为如前所述的多项式, 但是它们关于新基 $(x^2, x, 1)$ 的分量是前面所得分量的逆序排列, 因此 D 的新矩阵表示为

$$\begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

接下来我们求 D 的第三个矩阵表示, 取新的有序基

$$(1, 1+x, 1+x+x^2, 1+x+x^2+x^3)(V \text{ 的}), \quad (1, x, x^2)(W \text{ 的})$$

此时, 基元素进行如下变换:

$$D(1) = 0, \quad D(1+x) = 1, \quad D(1+x+x^2) = 1+2x,$$
$$D(1+x+x^2+x^3) = 1+2x+3x^2,$$

所以此时 D 的矩阵表示为

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

4.11　对角形矩阵表示的构造

由于当选定的基不同时, 给定线性变换的矩阵表示也不同, 所以我们自然希望能选出使得所得矩阵具有某种特别简单形式的基, 这样一来所得的表示原象与象之间关系的线性方程组的形式也比较简单. 下面的定理表明我们可以得到除了从左上角到右下角的对角线上的元素外, 其余所有元素都为零的矩阵, 这样的对角线称为**主对角线**(main diagonal). 更进一步, 我们可使主对角线的前面若干元均为 1, 其余元均为零, 且为 1 的元的个数等于该变换的秩. 我们称当 $i \neq k$ 时所有元 $t_{ik} = 0$ 的矩阵为**对角矩阵**.

定理 4.14　设 V 和 W 为有限维线性空间且 $\dim V = n$, $\dim W = m$, 假定 $T \in \mathscr{L}(V, W)$ 并令 $r = \dim T(V)$ 表示 T 的秩, 那么存在 V 的基 $(v_1, \cdots, v_n)$ 和 W 的基 $(w_1, \cdots, w_m)$ 使得

$$T(v_i) = w_i, \quad i = 1, 2, \cdots, r, \tag{4.14}$$

$$T(v_k) = O, \quad i = r+1, 2, \cdots, n. \tag{4.15}$$

因此, T 关于这对基的矩阵表示 (t_{ik}) 除 r 对角元

$$t_{11}=t_{22}=\cdots=t_{rr}=1$$

外的所有元都为零.

证明 首先我们构造 W 的一组基. 由于 $T(V)$ 是 W 的子空间且它的维数等于 r, 所以它的基包含 W 中的 r 个元素, 设它的基为 $w_1,\cdots,w_r$. 由定理 4.7, 这些元素构成 W 的某组基的子集, 因此存在 W 中元素 $w_{r+1},\cdots,w_m$, 使得

$$(w_1,\cdots,w_r,w_{r+1},\cdots,w_m) \tag{4.16}$$

为 W 的基.

接下来我们构造 V 的基. 基 (4.16) 的前 r 个元素 w_i 中的任意一个都是 V 中至少一个元素的象, 对 w_i, 任取 V 中一个这样的元素并令其为 v_i, 则对 $i=1,2,\cdots,r$, 都有 $T(v_i)=w_i$, 所以它们满足条件 (4.14). 再令 k 为零空间 $N(T)$ 的维数, 由零化度加秩定理 (定理 4.3), 我们有 $k+r=n$. 由于零空间的维数为 k, 所以它的基包含 k 个 V 中元素, 设它的基为 $v_{r+1},\cdots,v_{r+k}$, 这些元素满足条件 (4.15). 因此为了证明本定理, 我们必须证明有序集

$$(v_1,\cdots,v_r,v_{r+1},\cdots,v_{r+k}) \tag{4.17}$$

是 V 的一组基. 因为 $\dim V=n=r+k$, 所以只需证明这些元素线性无关.

假定这些元素的某个线性组合为零, 设

$$\sum_{i=1}^{r+k}c_iv_i=O. \tag{4.18}$$

将变换 T 作用在上式两边, 再由条件 (4.14) 和 (4.15), 我们得到

$$\sum_{i=1}^{r+k}c_iT(v_i)=\sum_{i=1}^{r}c_iw_i=O.$$

又因为 $w_1,\cdots,w_r$ 线性无关, 所以 $c_1=\cdots=c_r=0$, 因此式 (4.18) 中前 r 项为零, 于是可将它简化为

$$\sum_{i=r+1}^{r+k}c_iv_i=O.$$

由于 $v_{r+1},\cdots,v_{r+k}$ 构成 T 的零空间的基, 所以它们线性无关, 因此 $c_{r+1}=\cdots=c_{r+k}=0$. 于是式 (4.18) 中的所有系数 c_i 都为零, 从而有序集 (4.17) 中的元素构成 V 的一组基. 定理得证. □

例 参考 4.10 节中的例 6, 其中 D 为微分算子, 它将次数 $\leqslant 3$ 的多项式空间 V 映到次数 $\leqslant 2$ 的多项式空间 W 中. 在此例中, T 的值域为 W, 所以 T 的秩为 3. 应用定理 4.14 的证明中的方法, 我们任选 W 的一组基, 例如 $(1,x,x^2)$, 那么 D 将元素 $(x,\frac{1}{2}x^2,\frac{1}{3}x^3)$ 分别映为 $(1,x,x^2)$. 另外我们选常数多项式 1 作为 D 的零空间

的基并将它与 $(x, \frac{1}{2}x^2, \frac{1}{3}x^3)$ 合并作为 V 的基, 因此我们选 $(x, \frac{1}{2}x^2, \frac{1}{3}x^3, 1)$ 作为 V 的基, $(1, x, x^2)$ 作为 W 的基, 此时对应的 D 的矩阵具有如下对角形式:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

□

4.12 习　　题

本节中除非特别指明, 否则对所有的 $\mathbb{R}^n$, 我们都取单位坐标向量作为它们的基. 此外除非特别指明, 所有牵涉到线性变换 $T: V \to W$ 的矩阵表示的习题中, 都有 $V = W$, 并且 V 和 W 取相同的基.

1. 求 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 的以下各线性变换的矩阵表示.

   (a) 恒等变换.　(b) 零变换.　(c) 与给定纯量 c 的乘积.

2. 求以下各投影的矩阵.

   (a) $T: \mathbb{R}^3 \to \mathbb{R}^2$, 其中 $T(x_1, x_2, x_3) = (x_1, x_2)$.

   (b) $T: \mathbb{R}^3 \to \mathbb{R}^2$, 其中 $T(x_1, x_2, x_3) = (x_2, x_3)$.

   (c) $T: \mathbb{R}^5 \to \mathbb{R}^3$, 其中 $T(x_1, x_2, x_3, x_4, x_5) = (x_2, x_3, x_4)$.

3. 线性变换 $T: \mathbb{R}^2 \to \mathbb{R}^2$ 对基向量 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 进行如下映射:

$$T(\boldsymbol{i}) = \boldsymbol{i} + \boldsymbol{j}, \quad T(\boldsymbol{j}) = 2\boldsymbol{i} - \boldsymbol{j}$$

   (a) 用 $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 表示 $T(3\boldsymbol{i} - 4\boldsymbol{j})$ 和 $T^2(3\boldsymbol{i} - 4\boldsymbol{j})$.

   (b) 求 T 和 T^2 的矩阵表示.

   (c) 如果用 (e_1, e_2) 取代 $(\boldsymbol{i}, \boldsymbol{j})$ 作为 $\mathbb{R}^2$ 的基, 其中 $e_1 = \boldsymbol{i} - \boldsymbol{j}$, $e_2 = 3\boldsymbol{i} + \boldsymbol{j}$, 求 T 和 T^2 的矩阵表示.

4. 定义线性变换 $T: \mathbb{R}^2 \to \mathbb{R}^2$ 如下: 对每一个向量 (x, y) 都先做关于 y 轴的反射, 再使其长度变为原来的两倍从而得到 $T(x, y)$, 求 T 和 T^2 的矩阵.

5. 令 $T: \mathbb{R}^3 \to \mathbb{R}^3$ 为线性变换, 使得

$$T(\boldsymbol{k}) = 2\boldsymbol{i} + 3\boldsymbol{j} + 5\boldsymbol{k}, \quad T(\boldsymbol{j} + \boldsymbol{k}) = \boldsymbol{i}, \quad T(\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) = \boldsymbol{j} - \boldsymbol{k}.$$

   (a) 计算 $T(\boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k})$ 并求 T 的零化度和秩.

   (b) 求 T 的矩阵.

6. 对习题 5 中的变换, 设定义域和值域的基都为 (e_1, e_2, e_3), 其中 $e_1 = (2, 3, 5)$, $e_2 = (1, 0, 0)$, $e_3 = (0, 1, -1)$, 求 T 关于新基的矩阵表示.

7. 线性变换 $T: \mathbb{R}^3 \to \mathbb{R}^2$ 对基向量进行如下映射: $T(\boldsymbol{i}) = (0, 0)$, $T(\boldsymbol{j}) = (1, 1)$, $T(\boldsymbol{k}) = (1, -1)$.

   (a) 计算 $T(4\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k})$ 并求 T 的零化度和秩.

   (b) 求 T 的矩阵.

   (c) 令 $(\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k})$ 为 $\mathbb{R}^3$ 的基, (w_1, w_2) 为 $\mathbb{R}^2$ 的基, 其中 $w_1 = (1, 1)$, $w_2 = (1, 2)$, 求 T 关于新基的矩阵.

   (d) 求 $\mathbb{R}^3$ 的基 (e_1, e_2, e_3) 和 $\mathbb{R}^2$ 的基 (w_1, w_2) 使得 T 关于它们的矩阵为对角矩阵.

8. 线性变换 $T: \mathbb{R}^2 \to \mathbb{R}^3$ 对基向量进行如下映射: $T(\boldsymbol{i}) = (1, 0, 1)$, $T(\boldsymbol{j}) = (-1, 0, 1)$.

   (a) 计算 $T(2\boldsymbol{i} - 3\boldsymbol{j})$ 并求 T 的零化度和秩.

   (b) 求 T 的矩阵.

   (c) 求 $\mathbb{R}^2$ 的基 (e_1, e_2) 和 $\mathbb{R}^3$ 的基 (w_1, w_2, w_3) 使得 T 关于它们的矩阵为对角矩阵.

9. 若 $T(\boldsymbol{i}) = (1, 0, 1)$, $T(\boldsymbol{j}) = (1, 1, 1)$, 求解习题 8.

10. 设 V 和 W 为线性空间, 它们的维数都是 2 且基都为 (e_1, e_2), 令 $T: V \to W$ 为使 $T(e_1 + e_2) = 3e_1 + 9e_2$, $T(3e_1 + 2e_2) = 7e_1 + 23e_2$ 的线性变换.

(a) 计算 $T(e_2-e_1)$ 并求 T 的零化度和秩.

(b) 求 T 关于给定基的矩阵.

(c) 令 V 的基仍为 (e_1,e_2), 求 W 的形如 $(e_1+ae_2, 2e_1+be_2)$ 的基, 使得 T 关于这个新基的矩阵为对角矩阵.

习题 11~20 需要微积分知识.

在由全体实值函数组成的线性空间中, 下述各元素组都线性无关且生成一个有限维空间 V. 在各题中, 分别取给定元素组为 V 的基并令 $D: V\to V$ 为微分算子, 求 D 和 D^2 关于这样选取的基的矩阵.

11. $(\sin x, \cos x)$.
12. $(1, x, e^x)$.
13. $(1, 1+x, 1+x+e^x)$.
14. (e^x, xe^x).
15. $(-\cos x, \sin x)$.
16. $(\sin x, \cos x, x\sin x, x\cos x)$.
17. $(e^x\sin x, e^x\cos x)$.
18. $(e^{2x}\sin 3x, e^{2x}\cos 3x)$.
19. 设 V 为次数 $\leqslant 3$ 的所有实多项式组成的线性空间, 取 $(1,x,x^2,x^3)$ 为 V 的基, 令 D 为微分算子并令 $T: V\to V$ 为将 $p(x)$ 映为 $xp'(x)$ 的线性变换, 分别求以下各变换关于给定基的矩阵:

    (a) T. (b) DT. (c) TD.

    (d) $TD-DT$. (e) T^2. (f) $T^2D^2-D^2T^2$.
20. 参见习题 19, 令 W 为 V 在 TD 下的象, 求 V 和 W 的基, 使得 TD 关于它们的矩阵为对角矩阵.

4.13 矩阵组成的线性空间

如前几节所述, 矩阵是作为线性变换的表示而自然引入的, 然而矩阵本身也有不依赖于线性变换而独立存在的理由. 此时, 它们组成一类新的可以定义代数运算的数学对象. 与线性变换之间的关系是我们引入矩阵的动机, 但是我们现在暂时不考虑这种关系.

令 m 和 n 为两个正整数, 并令 $I_{m,n}$ 为所有整数对 (i,j) 组成的集合, 其中 $1\leqslant i\leqslant m$, $1\leqslant j\leqslant n$, 我们称定义域为 $I_{m,n}$ 的任意函数 $\boldsymbol{A}$ 为 $m\times n$ 矩阵, 称函数值 $\boldsymbol{A}(i,j)$ 为该矩阵的 **ij 元**或 **ij 元素**, 并将它记作 a_{ij}. 习惯上我们用如下 m 行 n 列矩形阵列表示它的所有函数值:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

元素 a_{ij} 可为任意一种对象. 通常它们是实数或复数, 然而有时使用元素为函数等其他对象的矩阵更为方便. 我们也用更紧凑的记号

$$\boldsymbol{A}=(a_{ik})_{i,k=1}^{m,n} \quad 或 \quad \boldsymbol{A}=(a_{ik}).$$

表示矩阵. 如果 $m=n$, 我们称该矩阵为**方阵** (square matrix), 我们也称 $1\times n$ 矩阵为**行矩阵**, 称 $m\times 1$ 矩阵为**列矩阵**.

当且仅当两个函数的定义域相同且定义域中每个元素的函数值都相等时, 这两

个函数相等. 由于矩阵是函数, 因此当且仅当两个矩阵 $\boldsymbol{A}=(a_{ij})$ 和 $\boldsymbol{B}=(b_{ij})$ 的行数和列数分别相等且对所有数对 (i,j) 都有 $a_{ij}=b_{ij}$ 时, 这两个矩阵相等.

现在我们假定矩阵的元素都是数 (可以是实数也可以是复数), 我们用与定义所有实值函数或复值函数的加法和纯量乘法相同的方法来定义矩阵的加法和纯量乘法.

定义 设 $\boldsymbol{A}=(a_{ij})$ 和 $\boldsymbol{B}=(b_{ij})$ 为两个 $m\times n$ 矩阵, c 为任意纯量, 定义矩阵 $\boldsymbol{A}+\boldsymbol{B}$ 和 $c\boldsymbol{A}$ 如下:

$$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij}),\quad c\boldsymbol{A}=(ca_{ij}).$$

仅当 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的行数和列数分别相等时, 它们的和才有意义.

例 若

$$\boldsymbol{A}=\begin{bmatrix}1 & 2 & -3\\ -1 & 0 & 4\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}5 & 0 & 1\\ 1 & -2 & 3\end{bmatrix},$$

那么我们有

$$\boldsymbol{A}+\boldsymbol{B}=\begin{bmatrix}6 & 2 & -2\\ 0 & -2 & 7\end{bmatrix},\quad 2\boldsymbol{A}=\begin{bmatrix}2 & 4 & -6\\ -2 & 0 & 8\end{bmatrix},$$

$$(-1)\boldsymbol{B}=\begin{bmatrix}-5 & 0 & -1\\ -1 & 2 & -3\end{bmatrix}.$$ □

我们将元素全为零的 $m\times n$ 矩阵定义为零矩阵 $\boldsymbol{O}$. 由这些定义易证所有 $m\times n$ 矩阵组成的集合是一个线性空间, 我们将这个线性空间记作 $M_{m,n}$. 当矩阵的元素为实数时, $M_{m,n}$ 是实线性空间; 当它们的元素为复数时, $M_{m,n}$ 是一个复线性空间. 易证 $M_{m,n}$ 的维数为 mn. 事实上, 只有一个元为 1, 其余元为 0 的 mn 个不同矩阵构成 $M_{m,n}$ 的一组基. 例如, 下面六个矩阵

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},\quad \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix},\quad \begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix},$$

$$\begin{bmatrix}0 & 0 & 0\\ 1 & 0 & 0\end{bmatrix},\quad \begin{bmatrix}0 & 0 & 0\\ 0 & 1 & 0\end{bmatrix},\quad \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 1\end{bmatrix},$$

构成所有 2×3 矩阵组成的线性空间的一组基.

4.14 线性变换与矩阵之间的同构

我们回过头来考虑矩阵和线性变换之间的关系. 令 V 和 W 为有限维线性空间且 $\dim V=n, \dim W=m$, 选定 V 的基 $(v_1,\cdots,v_n)$ 和 W 的基 $(w_1,\cdots,w_m)$, 在接下来的讨论中, 我们将一直选定这两组基. 令 $\mathscr{L}(V,W)$ 表示 V 到 W 的所有线性变换组成的线性空间. 若 $T\in\mathscr{L}(V,W)$, 则令 $m(T)$ 表示 T 关于上述两组基的矩

阵表示. 由前面的讨论可知, $m(T)$ 可如下确定.

将每一个基元素 v_k 的象都表示为 W 中基元素的线性组合:

$$T(v_k) = \sum_{i=1}^{m} t_{ik} w_i, \quad k = 1, 2, \cdots, n. \tag{4.19}$$

则 t_{ik} 就是 $m(T)$ 的 ik 元, 因此我们有

$$m(T) = (t_{ik})_{i,k=1}^{m,n}. \tag{4.20}$$

等式 (4.20) 定义了一个新函数 m, 它的定义域为 $\mathscr{L}(V, W)$, 它的值都是 $M_{m,n}$ 中的矩阵. 由于对每个 $m \times n$ 矩阵, 都存在 $\mathscr{L}(V, W)$ 中的 T 使它等于 $m(T)$, 所以 m 的值域为 $M_{m,n}$. 下面的定理告诉我们, 变换 $m : \mathscr{L}(V, W) \to M_{m,n}$ 是定义在 $\mathscr{L}(V, W)$ 上的一一线性变换.

定理 4.15 (同构定理) *对 $\mathscr{L}(V, W)$ 中任意 S 和 T 以及任意纯量 c, 我们有*

$$m(S + T) = m(S) + m(T), \quad m(cT) = cm(T).$$

更进一步:

$$m(S) = m(T) \quad \text{必然包含} \quad S = T,$$

所以 m 在 $\mathscr{L}(V, W)$ 上一一.

证明 矩阵 $m(T)$ 由表达式 (4.19) 中的系数 t_{ik} 组成, 类似地, 矩阵 $m(S)$ 由下述各式

$$S(v_k) = \sum_{i=1}^{m} s_{ik} w_i, \quad k = 1, 2, \cdots, n. \tag{4.21}$$

中的系数 s_{ik} 组成. 又因为我们有

$$(S + T)(v_k) = \sum_{i=1}^{m} (s_{ik} + t_{ik}) w_i, \quad (cT)(v_k) = \sum_{i=1}^{m} (ct_{ik}) w_i,$$

所以可知 $m(S + T) = (s_{ik} + t_{ik}) = m(S) + m(T)$ 以及 $m(cT) = (ct_{ik}) = c(t_{ik}) = cm(T)$. 于是 m 的线性性得证.

为了证明 m 在 $\mathscr{L}(V, W)$ 上一一, 假定 $m(S) = m(T)$, 其中 $S = (s_{ik})$, $T = (t_{ik})$, 等式 (4.19) 和 (4.21) 表明对任意基元素 v_k, 都有 $S(v_k) = T(v_k)$, 所以对任意 V 中的 x, 都有 $S(x) = T(x)$, 因此 $S = T$. □

注 函数 m 称为一个同构. 对给定的一组基, m 建立了所有线性变换组成的集合 $\mathscr{L}(V, W)$ 与所有 $m \times n$ 矩阵组成的集合 $M_{m,n}$ 之间的一个一一对应关系, 加法和纯量乘法在这个对应关系下保持不变. 我们称线性空间 $\mathscr{L}(V, W)$ 和 $M_{m,n}$ **同构**(isomorphic). 顺便提一下, 定理 4.11 表明一一线性变换的定义域和值域的维数相同, 因此 $\dim \mathscr{L}(V, W) = \dim M_{m,n} = mn$.

如果 $V = W$ 且我们为 V 和 W 选的基相同, 那么恒等变换 $I : V \to V$ 对应的矩阵 $m(I)$ 是一个 $n \times n$ 对角矩阵, 其主对角线上的元都为 1, 其余元都为 0. 我们称这样的矩阵为**恒等矩阵** (identity matrix) 或**单位矩阵** (unit matrix), 并将它记作 $\boldsymbol{I}$, 当我们想要强调它是一个 $n \times n$ 矩阵时, 将它记作 $\boldsymbol{I}_n$.

4.15 矩阵的乘法

有些线性变换可通过复合来进行相乘. 接下来我们将定义矩阵的乘法, 在这种定义下, 两个矩阵的乘积对应它们各自代表的线性变换的复合.

通过前几节的讨论我们知道, 如果 $T: U \to V$ 和 $S: V \to W$ 是两个线性变换, 那么它们的复合 $ST: U \to W$ 是由

$$(ST)(x) = S[T(x)], \quad 对 U 中任何 x$$

确定的线性变换.

假定 U, V, W 都是有限维空间, 设

$$\dim U = n, \quad \dim V = p, \quad \dim W = m.$$

分别选定 U, V, W 的基. 关于这些基的矩阵表示 $m(S)$ 是一个 $m \times p$ 矩阵, 矩阵表示 $m(T)$ 是一个 $p \times n$ 矩阵, 矩阵表示 $m(ST)$ 是一个 $m \times n$ 矩阵. 通过下面的矩阵乘法的定义, 我们可以推导关系 $m(ST) = m(S)m(T)$ 并从而将同构关系推广到乘积上.

定义 (矩阵的乘积) 令 $\boldsymbol{A}$ 为任意 $m \times p$ 矩阵, $\boldsymbol{B}$ 为任意 $p \times n$ 矩阵, 设

$$\boldsymbol{A} = (a_{ij})_{i,j=1}^{m,p}, \quad \boldsymbol{B} = (b_{ij})_{i,j=1}^{p,n}.$$

则定义 (有序) 积 $\boldsymbol{AB}$ 为 $m \times n$ 矩阵 $\boldsymbol{C} = (c_{ij})$, 其中 $\boldsymbol{C}$ 的 ij 元素由

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} \tag{4.22}$$

确定.

注 除非左因子 $\boldsymbol{A}$ 的列数与右因子 $\boldsymbol{B}$ 的行数相等, 否则积 $\boldsymbol{AB}$ 无意义.

在本节中, 我们用记号 $\boldsymbol{A}_i$ 表示 $\boldsymbol{A}$ 的第 i 行, $\boldsymbol{B}^j$ 表示 $\boldsymbol{B}$ 的第 j 列. 若将它们看作 p 维向量, 则式 (4.22) 中的和即为点积 $\boldsymbol{A}_i \cdot \boldsymbol{B}^j$, 也就是说, $\boldsymbol{AB}$ 的 ij 元素是 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 j 列的点积:

$$\boldsymbol{AB}(\boldsymbol{A}_i \cdot \boldsymbol{B}^j)_{ij=1}^{m,n}.$$

因此我们可将矩阵乘法看作是点积的推广.

例 1 令 $\boldsymbol{A} = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 0 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 4 & 6 \\ 5 & -1 \\ 0 & 2 \end{bmatrix}$, 由于 $\boldsymbol{A}$ 是一个 2×3 矩阵, $\boldsymbol{B}$ 是一个 3×2 矩阵, 所以 $\boldsymbol{AB}$ 是一个 2×2 矩阵. 将 $\boldsymbol{A}$ 的各行看作是 $\mathbb{R}^3$ 中的向量, 那么我们有 $\boldsymbol{A}_1 = (3, 1, 2)$, $\boldsymbol{A}_2 = (-1, 1, 0)$, 将 $\boldsymbol{B}$ 的各列看作是 $\mathbb{R}^3$ 中向量, 那么我们有 $\boldsymbol{B}^1 = (4, 5, 0)$, $\boldsymbol{B}^2 = (6, -1, 2)$, 因此积 $\boldsymbol{AB}$ 由

$$\boldsymbol{AB} \begin{bmatrix} \boldsymbol{A}_1 \cdot \boldsymbol{B}^1 & \boldsymbol{A}_1 \cdot \boldsymbol{B}^2 \\ \boldsymbol{A}_2 \cdot \boldsymbol{B}^1 & \boldsymbol{A}_2 \cdot \boldsymbol{B}^2 \end{bmatrix} = \begin{bmatrix} 17 & 21 \\ 1 & -7 \end{bmatrix}.$$

确定. □

例 2 令 $\boldsymbol{A}=\begin{bmatrix}2&1&-3\\1&2&4\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}-2\\1\\2\end{bmatrix}$, 则 $\boldsymbol{A}$ 是一个 2×3 矩阵, $\boldsymbol{B}$ 是一个 3×1 矩阵, 所以 $\boldsymbol{AB}$ 是由

$$\boldsymbol{AB}=\begin{bmatrix}\boldsymbol{A}_1\cdot\boldsymbol{B}^1\\\boldsymbol{A}_2\cdot\boldsymbol{B}^1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}-9\\8\end{bmatrix}.$$

确定的 2×1 矩阵. □

例 3 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为大小相同的方阵, 那么 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 都有意义. 例如, 若

$$\boldsymbol{A}=\begin{bmatrix}1&2\\-1&1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}3&4\\5&2\end{bmatrix},$$

那么我们可得

$$\boldsymbol{AB}=\begin{bmatrix}13&8\\2&-2\end{bmatrix},\quad \boldsymbol{BA}=\begin{bmatrix}-1&10\\3&12\end{bmatrix}.$$

本例表明, 一般来说, $\boldsymbol{AB}\neq\boldsymbol{BA}$. 如果 $\boldsymbol{AB}=\boldsymbol{BA}$, 那么我们说 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 可交换.

例 4 设 $\boldsymbol{I}_p$ 为 $p\times p$ 单位矩阵, 则对任意 $p\times n$ 矩阵 $\boldsymbol{A}$ 都有 $\boldsymbol{I}_p\boldsymbol{A}=\boldsymbol{A}$, 对任意 $m\times p$ 矩阵 $\boldsymbol{B}$ 都有 $\boldsymbol{BI}_p=\boldsymbol{B}$. 例如

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}2\\3\\4\end{bmatrix}=\begin{bmatrix}2\\3\\4\end{bmatrix},\quad \begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}.$$ □

接下来我们证明复合变换 ST 的矩阵表示等于 S 的矩阵表示 $m(S)$ 与 T 的矩阵表示 $m(T)$ 之积.

定理 4.16 设 $T:U\to V$ 和 $S:V\to W$ 为两个线性变换, 其中 U,V,W 为有限维线性空间, 那么在选定 U,V,W 的基之后, S,T,ST 的矩阵表示之间的关系满足等式

$$m(ST)=m(S)m(T).$$

证明 假定 $\dim U=n$, $\dim V=p$, $\dim W=m$, 分别选定线性空间 U,V,W 的基, 设 U 的基为 $(u_1,\cdots,u_n)$, V 的基为 $(v_1,\cdots,v_p)$, W 的基为 $(w_1,\cdots,w_m)$, 那么 S 和 T 关于这些基的矩阵表示分别为

$$m(S)=(s_{ij})_{i,j=1}^{m,p},\quad 其中\quad S(v_k)=\sum_{i=1}^{m}s_{ik}w_i,\quad k=1,2,\cdots,p,$$

$$m(T)=(t_{ij})_{ij=1}^{p,n},\quad 其中\quad T(u_j)=\sum_{k=1}^{p}t_{kj}v_k,\quad j=1,2,\cdots,n.$$

因此我们有

$$ST(u_j) = S[T(u_j)] = \sum_{k=1}^{p} t_{kj} S(v_k) = \sum_{k=1}^{p} t_{kj} \sum_{i=1}^{m} s_{ik} w_i = \sum_{i=1}^{m} \left(\sum_{k=1}^{p} s_{ik} t_{kj} \right) w_i,$$

于是我们有

$$m(ST) = \left(\sum_{k=1}^{p} s_{ik} t_{kj} \right)_{i,j=1}^{m,n} = m(S)m(T). \qquad \square$$

我们已经指出矩阵的乘法并不总是可交换的. 下面的定理表明它总是满足结合律和分配律.

定理 4.17 (矩阵乘法的结合律和分配律) 给定矩阵 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$.

(a) 如果积 $\boldsymbol{A}(\boldsymbol{BC})$ 和 $(\boldsymbol{AB})\boldsymbol{C}$ 都有意义, 那么我们有

$$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C} \quad (\text{结合律}).$$

(b) 假定 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的行数相同, 列数也相同, 如果 $\boldsymbol{AC}$ 和 $\boldsymbol{BC}$ 都有意义, 那么我们有

$$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC} \quad (\text{右分配律}).$$

此外如果 $c\boldsymbol{A}$ 和 $c\boldsymbol{B}$ 都有意义, 那么我们有

$$\boldsymbol{C}(\boldsymbol{A} + \boldsymbol{B}) = \boldsymbol{CA} + \boldsymbol{CB} \quad (\text{左分配律}).$$

证明 可由矩阵乘法的定义直接推导这几条性质, 但是我们更希望用线性变换与矩阵之间的联系来证明本定理. 引入有限维线性空间 U, V, W, X 和定义在它们上面的三个线性变换 $T: U \to V$, $S: V \to W$, $R: W \to X$, 使得对选定的基, 我们有

$$\boldsymbol{A} = m(R), \quad \boldsymbol{B} = m(S), \quad \boldsymbol{C} = m(T).$$

由定理 4.16, 我们有 $m(RS) = \boldsymbol{AB}$ 和 $m(ST) = \boldsymbol{BC}$. 由函数复合的结合律, 我们可知 $R(ST) = (RS)T$, 对此式再一次用定理 4.16, 我们得到 $m(R)m(ST) = m(RS)m(T)$, 即 $\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C}$, 这证明了性质 (a), 同理可证性质 (b). $\square$

定义 (矩阵的幂) 若 $\boldsymbol{A}$ 是一个方阵, 则归纳定义 $\boldsymbol{A}$ 的整数幂如下:

$$\boldsymbol{A}^0 = \boldsymbol{I}, \quad \boldsymbol{A}^n = \boldsymbol{A}\boldsymbol{A}^{n-1}, \quad n \geqslant 1.$$

读者可证明同一个矩阵的整数幂可交换, 即 $\boldsymbol{A}^n\boldsymbol{A}^m = \boldsymbol{A}^m\boldsymbol{A}^n$, 且它满足指数定律 $\boldsymbol{A}^n\boldsymbol{A}^m = \boldsymbol{A}^{n+m}$.

4.16 习 题

1. 给定矩阵 $\boldsymbol{A} = \begin{bmatrix} 1 & -4 & 2 \\ -1 & 4 & -2 \end{bmatrix}$, $\boldsymbol{B} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 5 & -2 \end{bmatrix}$, $\boldsymbol{C} = \begin{bmatrix} 2 & 2 \\ 1 & -1 \\ 1 & -3 \end{bmatrix}$, 计算 $\boldsymbol{B}+\boldsymbol{C}$, $\boldsymbol{AB}$, $\boldsymbol{BA}$, $\boldsymbol{AC}$, $\boldsymbol{CA}$ 和 $\boldsymbol{A}(2\boldsymbol{B} - 3\boldsymbol{C})$.

2. 令 $A=\begin{bmatrix}0&1\\0&2\end{bmatrix}$, 求所有 2×2 矩阵 B, 使得 (a) $AB=O$, (b) $BA=O$.

3. 分别为以下各题求满足给定等式的 a,b,c,d:

(a) $\begin{bmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\end{bmatrix}=\begin{bmatrix}1\\9\\6\\5\end{bmatrix}$.

(b) $\begin{bmatrix}a&b&c&d\\1&4&9&2\end{bmatrix}\begin{bmatrix}1&0&2&0\\0&0&1&1\\0&1&0&0\\0&0&1&0\end{bmatrix}=\begin{bmatrix}1&0&6&6\\1&9&8&4\end{bmatrix}$.

4. 为以下各矩阵 A 和 B 分别计算 $AB-BA$:

(a) $A=\begin{bmatrix}1&2&2\\2&1&2\\1&2&3\end{bmatrix}$, $B=\begin{bmatrix}4&1&1\\-4&2&0\\1&2&1\end{bmatrix}$.

(b) $A=\begin{bmatrix}2&0&0\\1&1&2\\-1&2&1\end{bmatrix}$, $B=\begin{bmatrix}3&1&-2\\3&-2&4\\-3&5&11\end{bmatrix}$.

5. 如果 A 是方阵, 证明对所有整数 $m\geqslant 0$, $n\geqslant 0$, 都有 $A^nA^m=A^mA^n$.

6. 令 $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$, 验证 $A^2=\begin{bmatrix}1&2\\0&1\end{bmatrix}$ 并计算 A^n.

7. 令 $A=\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}$, 验证 $A^2=\begin{bmatrix}\cos 2\theta&-\sin 2\theta\\\sin 2\theta&\cos 2\theta\end{bmatrix}$ 并计算 A^n.

8. 令 $A=\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}$, 验证 $A^2=\begin{bmatrix}1&2&3\\0&1&2\\0&0&1\end{bmatrix}$, 计算 A^3 和 A^4, 猜测一个计算 A^n 的公式并用归纳法证明之.

9. 令 $A=\begin{bmatrix}1&0\\-1&1\end{bmatrix}$, 证明 $A^2=2A-I$ 并计算 A^{100}.

10. 求满足 $A^2=O$ 的所有 2×2 矩阵 A.

11. (a) 证明当且仅当 2×2 矩阵 A 与四个矩阵

$$\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad\begin{bmatrix}0&1\\0&0\end{bmatrix},\quad\begin{bmatrix}0&0\\1&0\end{bmatrix},\quad\begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

都可交换时, A 与所有 2×2 矩阵都可交换.

(b) 求所有这样的矩阵 A.

12. 2×2 矩阵

$$\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad\begin{bmatrix}1&0\\c&-1\end{bmatrix},\quad\begin{bmatrix}1&b\\0&-1\end{bmatrix},$$

都满足等式 $A^2=I$, 其中 b 和 c 为任意实数. 求所有满足 $A^2=I$ 的矩阵 A.

13. 若

$$A=\begin{bmatrix}2&-1\\-2&3\end{bmatrix},\quad B=\begin{bmatrix}7&6\\9&8\end{bmatrix}.$$

求所有 2×2 矩阵 $\boldsymbol{C}$ 和 D, 使得 $\boldsymbol{AC}=\boldsymbol{B}$ 以及 $D\boldsymbol{A}=\boldsymbol{B}$.

14. (a) 验证 2×2 矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & -1\\ 0 & 2\end{bmatrix}$ 和 $\boldsymbol{B}=\begin{bmatrix}1 & 0\\ 1 & 2\end{bmatrix}$ 不满足代数公式

$$(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2,\quad (\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2$$

(b) 修改这些公式的右边使所得公式对所有方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都成立.

(c) 什么样的矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 满足 (a) 中的公式?

15. 为以下关于 $n\times n$ 矩阵的各命题提供证明或举出反例:

(a) 若 $(\boldsymbol{A}-\boldsymbol{B})^2=(\boldsymbol{A}+\boldsymbol{B})^2$, 则 $\boldsymbol{A}^2\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}^2$.

(b) 若 $\boldsymbol{A}^2=\boldsymbol{I}$, 则 $\boldsymbol{A}=\boldsymbol{I}$ 或 $\boldsymbol{A}=-\boldsymbol{I}$.

4.17 在线性方程组中的应用

令 $\boldsymbol{A}=(a_{ij})$ 为给定 $m\times n$ 矩阵, 其中 a_{ij} 为数, 并令 $c_1,\cdots,c_m$ 为 m 个数, 那么形如

$$\sum_{k=1}^{n}a_{ik}x_k=c_i,\quad i=1,2,\cdots,m,\tag{4.23}$$

的方程组称为由 m 个 n 元线性方程组成的方程组. 这里我们将 $x_1,\cdots,x_n$ 看作未知元. 我们称满足方程组中所有方程的任意 n 元组为该方程组的一个**解**, 称矩阵 $\boldsymbol{A}=(a_{ij})$ 为该方程组的**系数矩阵**.

可用线性变换语言来描述线性方程组. 取单位坐标向量作为 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 的基, 那么系数矩阵 $\boldsymbol{A}$ 确定一个线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$, 该线性变换将 $\mathbb{R}^n$ 中向量 $\boldsymbol{x}=(x_1,\cdots,x_n)$ 映为 $\mathbb{R}^m$ 中由如下 m 个线性方程给出的向量 $(y_1,\cdots,y_m)$:

$$y_i=\sum_{k=1}^{n}a_{ik}x_k,\quad i=1,2,\cdots,m.$$

方程组 (4.23) 可简写为

$$T(\boldsymbol{x})=\boldsymbol{c},$$

其中 $\boldsymbol{c}=(c_1,\cdots,c_m)$ 为 $\mathbb{R}^m$ 中的向量且它的各分量分别为式 (4.23) 中的各常数项. 当且仅当 c 在 T 的值域中时, 方程组有解. 若 $\mathbb{R}^n$ 中恰有一个 $\boldsymbol{x}$ 的象为 c, 那么方程组有唯一解, 若有多个 x 的象为 c, 那么方程组有多解.

例 1 (一个无解的方程组) 方程组 $x+y=1, x+y=2$ 由两个二元一次方程组成, 由于两个数之和不可能既是 1 又是 2, 所以该方程组无解. □

例 2 (有唯一解的方程组) 方程组 $x+y=1, x-y=0$ 由两个二元一次方程组成, 它有唯一解 $(x,y)=(\frac{1}{2},\frac{1}{2})$. □

例 3 (有多解的方程组) 方程组 $x+y=1$ 由一个二元一次方程组成, 它有无穷多解. 任意两个和为 1 的数都是它的解. □

对任意方程组 (4.23), 我们可将它与另一个方程组

$$\sum_{k=1}^{n} a_{ik}x_k = 0, \quad i = 1, 2, \cdots, m, \tag{4.24}$$

联系起来, 可用 0 代替方程组 (4.23) 中所有 c_i 从而得到这个新方程组. 我们称这个新方程组为对应于方程组 (4.23) 的**齐次线性方程组**. 如果 $\boldsymbol{c} \neq \boldsymbol{O}$, 则称方程组 (4.23) 为**非齐次线性方程组**. 当且仅当 $\mathbb{R}^n$ 中向量 $\boldsymbol{x}$ 满足

$$T(\boldsymbol{x}) = \boldsymbol{O},$$

时, 称它满足该齐次方程组, 其中 T 为由该系数矩阵确定的线性变换. 齐次方程组总有解 $\boldsymbol{x} = \boldsymbol{O}$, 它也可能有其他解. 齐次方程组的解集是 T 的零空间. 下面的定理描述了齐次方程组的解与对应的非齐次方程组的解之间的关系.

定理 4.18 *假定非齐次方程组 (4.23) 有解 $\boldsymbol{b}$.*

(a) *如果向量 $\boldsymbol{x}$ 是非齐次方程组 (4.23) 的解, 那么向量 $\boldsymbol{v} = \boldsymbol{x} - \boldsymbol{b}$ 是它对应的齐次方程组 (4.24) 的解.*

(b) *如果向量 $\boldsymbol{v}$ 是齐次方程组 (4.24) 的解, 那么向量 $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{b}$ 是非齐次方程组 (4.23) 的解.*

证明 令 $T : \mathbb{R}^n \to \mathbb{R}^m$ 为如前所述的系数矩阵 $\boldsymbol{A}$ 确定的线性变换, 由于 $\boldsymbol{b}$ 是非齐次方程组 (4.23) 的解, 所以我们有 $T(\boldsymbol{b}) = \boldsymbol{c}$, 令 $\boldsymbol{x}$ 和 $\boldsymbol{v}$ 为 $\mathbb{R}^n$ 中两个向量且 $\boldsymbol{v} = \boldsymbol{x} - \boldsymbol{b}$, 则由 T 的线性性我们有

$$T(\boldsymbol{v}) = T(\boldsymbol{x} - \boldsymbol{b}) = T(\boldsymbol{x}) - T(\boldsymbol{b}) = T(\boldsymbol{x}) - \boldsymbol{c}.$$

所以当且仅当 $T(\boldsymbol{v}) = \boldsymbol{O}$ 时, $T(\boldsymbol{x}) = \boldsymbol{c}$, 这同时证明了 (a) 和 (b). □

上述定理表明, 可将求非齐次方程组的所有解的问题很自然地分为两部分: (1) 求齐次方程组的所有解 $\boldsymbol{v}$(T 的零空间); (2) 求非齐次方程组的一个特解 $\boldsymbol{b}$. 再将 $\boldsymbol{b}$ 与 T 的零空间中的所有向量相加, 就得到非齐次方程组的所有解 $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{b}$.

令 k 表示 T 的零空间 $N(T)$ 的维数. 如果我们找到齐次方程组的 k 个线性无关解 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$, 那么它们将构成 $N(T)$ 的一组基, 而且我们可以通过构造所有可能的线性组合

$$\boldsymbol{v} = t_1\boldsymbol{v}_1 + \cdots + t_k\boldsymbol{v}_k,$$

获得 $N(T)$ 中的所有 $\boldsymbol{v}$, 其中 $t_1, \cdots, t_k$ 为纯量, 我们称该线性组合为齐次方程组的**通解**(general solution). 如果 $\boldsymbol{b}$ 是非齐次方程组的一个特解, 那么

$$\boldsymbol{x} = \boldsymbol{b} + t_1\boldsymbol{v}_1 + \cdots + t_k\boldsymbol{v}_k.$$

确定它的所有解 x, 我们称这个线性组合为非齐次方程组的通解. 这样我们可将定理 4.18 重新描述如下:

定理 4.19 *令 $T : \mathbb{R}^n \to \mathbb{R}^m$ 为满足 $T(\boldsymbol{x}) = \boldsymbol{y}$ 的线性变换, 其中 $\boldsymbol{x} = (x_1, \cdots, x_n)$, $\boldsymbol{y} = (y_1, \cdots, y_m)$, 且*

$$y_i = \sum_{k=1}^{n} a_{ik} x_k, \quad i = 1, 2, \cdots, m.$$

令 k 表示 T 的零化度, 如果 $v_1, \cdots, v_k$ 为齐次方程组 $T(x) = O$ 的 k 个线性无关解, 且 b 为非齐次方程组 $T(x) = c$ 的一个特解, 那么非齐次方程组的通解为

$$x = b + t_1 v_1 + \cdots + t_k v_k,$$

其中 $t_1, \cdots, t_k$ 为任意纯量. □

上述定理并没有说明如何判断非齐次方程组是否有特解 b, 它也没有说明如何求齐次方程组的解 $v_1, \cdots, v_k$. 然而它确实告诉我们当非齐次方程组有解时, 我们将会得到什么. 下面的例子尽管很简单, 但是它演示了这个定理.

例 包含一个二元方程 $x+y=2$ 的方程组对应的齐次方程组包含方程 $x+y=0$, 因此, 零空间包含 $\mathbb{R}^2$ 中形如 $(t, -t)$ 的所有向量, 其中 t 为纯量, 由于 $(t, -t) = t(1, -1)$, 所以它是 $\mathbb{R}^2$ 的一个一维子空间, 且 $(1, -1)$ 是它的基向量. 非齐次方程组的一个特解是 $(0, 2)$, 因此非齐次方程组的通解由

$$(x, y) = (0, 2) + t(1, -1), \quad \text{或} \quad x = t, \quad y = 2 - t,$$

确定, 其中 t 为任意纯量. □

4.18 计算技术 · Gauss-Jordan 消元法

接下来我们讨论如何具体计算非齐次线性方程组的解的问题. 尽管有很多方法可以用来解这个问题, 但是当方程组较大时, 它们的计算量都很大.

我们将讨论一个广泛使用的方法, 这个方法称为**Gauss-Jordan 消元法** (elimination method). 与其他方法相比, 它比较简单且比较容易在计算机上编写程序. 事实上, 人们为这种方法编写了很多求解特定大规模方程组的不同的计算机程序.

Gauss-Jordan 消元法包含三类作用在线性方程组的各个方程上的基本运算:

(1) 交换两个方程;

(2) 用一个非零纯量乘一个方程中的所有项;

(3) 将一个方程的纯量倍数加到另一个方程上.

每当对方程组进行一个这样的运算时, 我们都将得到一个新的方程组, 而且这两个方程组的解集相同. 我们称解集相同的两个方程组**等价**(equivalent). 对方程组系统地进行这样的运算后, 我们最终将得到一个等价的方程组, 而且通过观察即可对它求解.

我们将通过一些例子来演示这种方法, 了解了这些例子, 读者就会清楚在一般情况下如何使用这种方法求解方程组.

例 1 (有唯一解的方程组) 考虑方程组

$$\begin{cases} 2x-5y+4z=-3 \\ x-2y+z=5 \\ x-4y+6z=10. \end{cases}$$

这个方程组有一个特解 $x=124, y=75, z=31$, 我们将使用 Gauss-Jordan 消元法获得这个解. 为了简便起见, 我们并不打算不断重复书写 x, y, z 和等号, 而是对**增广矩阵**(augmented matrix)

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right] \tag{4.25}$$

作相应的运算, 这里增广矩阵是通过将方程组中等号右边的部分写在系数矩阵的右边而得到的矩阵. 前面所述的三种基本运算是作用在增广矩阵的行上的, 所以称作**行运算** (row operation). 在求解过程中的任意一步, 我们都可插入字母 x, y, z 并沿着竖线插入等号, 从而重新得到各方程. 我们的最终目的是在进行一系列运算后获得增广矩阵

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \end{array}\right] \tag{4.26}$$

它对应的方程组为

$$x=124, \quad y=75, \quad z=31,$$

它给出了我们需要的解.

求解过程的第一步是使矩阵的左上角的元素为 1. 为此, 我们可将矩阵 (4.25) 的第一行与第二行或第三行交换, 也可以用 $\frac{1}{2}$ 乘它的第一行, 这里我们选择交换矩阵的前两行, 此时我们有

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 2 & -5 & 4 & -3 \\ 1 & -4 & 6 & 10 \end{array}\right]$$

下一步是使第一列的其余元素为零且保持第一行不变. 为此我们用 -2 乘第一行并将所得结果加到第二行上, 然后再用 -1 乘第一行并将所得结果加到第三行, 通过这两个运算我们得到

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & -1 & 2 & -13 \\ 0 & -2 & 5 & 5 \end{array}\right]. \tag{4.27}$$

现在我们对与两个零相邻的较小的矩阵 $\left[\begin{array}{cc|c} -1 & 2 & -13 \\ -2 & 5 & 5 \end{array}\right]$ 重复上述过程. 可用 -1

乘矩阵 (4.27) 的第二行使它的左上角为 1, 这样得到矩阵

$$\left[\begin{array}{rrr|r} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & -2 & 5 & 5 \end{array}\right].$$

用 2 乘第二行并将所得结果加到第三行, 我们得到

$$\left[\begin{array}{rrr|r} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]. \tag{4.28}$$

此时对应的方程组为

$$\begin{cases} x - 2y + z = 5 \\ y - 2z = 13 \\ z = 31. \end{cases}$$

最后一个方程告诉我们 z 的值. 为了求 x 和 y, 我们由下到上求解方程, 从而得到

$$z = 31, \quad y = 13 + 2z = 13 + 62 = 75, \quad x = 5 + 2y - z = 5 + 150 - 31 = 124.$$

我们也可以继续进行 Gauss-Jordan 过程并将第二列和第三列中对角线上方的所有元化为零. 用 2 乘矩阵 (4.28) 的第二行并将所得结果加到第一行, 我们得到

$$\left[\begin{array}{rrr|r} 1 & 0 & -3 & 31 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]$$

最后, 我们用 3 乘第三行并将所得结果加到第一行, 然后再用 2 乘第三行并将所得结果加到第二行, 从而得到矩阵 (4.26). □

例 2 (有多个解的方程组) 考虑下面由三个五元方程组成的方程组:

$$\begin{cases} 2x - 5y + 4z + u - v = -3 \\ x - 2y + z - u + v = 5 \\ x - 4y + 6z - 2u - v = 10. \end{cases} \tag{4.29}$$

它对应的增广矩阵为

$$\left[\begin{array}{rrrrr|r} 2 & -5 & 4 & 1 & -1 & -3 \\ 1 & -2 & 1 & -1 & 1 & 5 \\ 1 & -4 & 6 & 2 & -1 & 10 \end{array}\right].$$

其中 x, y, z 的系数和等号右边的元素与例 1 中的相同. 当我们进行与例 1 中相同的行运算后, 最终将得到增广矩阵

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -16 & 19 & 124 \\ 0 & 1 & 0 & -9 & 11 & 75 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right].$$

可用 u 和 v 求解上式对应的方程组中的 x, y, z:

$$\begin{cases} x = 124 + 16u - 19v \\ y = 75 + 9u - 11v \\ z = 31 + 3u - 4v. \end{cases}$$

所以我们可用 u 和 v 表示 $\mathbb{R}^5$ 中的解向量 (x, y, z, u, v) 如下:

$$(x, y, z, u, v) = (124 + 16u - 19v, 75 + 9u - 11v, 31 + 3u - 4v, u, v).$$

合并 u 和 v 的同类项, 我们可重写解向量为如下形式:

$$(x, y, z, u, v) = (124, 75, 31, 0, 0) + u(16, 9, 3, 1, 0) + v(-19, -11, -4, 0, 1).$$

该等式用两个参数 u 和 v 给出了方程组的通解. 向量 $(124, 75, 31, 0, 0)$ 是非齐次方程组 (4.29) 的一个特解; 向量 $(16, 9, 3, 1, 0)$ 和向量 $(-19, -11, -4, 0, 1)$ 是对应的齐次方程组的两个解, 由于它们线性无关, 所以它们构成了齐次方程组的解空间的一组基. □

例 3 (无解的方程组) 考虑方程组

$$\begin{cases} 2x - 5y + 4z = -3 \\ x - 2y + z = 5 \\ x - 4y + 5z = 10. \end{cases} \tag{4.30}$$

该方程组与例 1 中的方程组几乎完全相同, 它们的差别仅在于第三个方程中 z 的系数: 它在例 1 中为 6, 在本例中为 5. 方程组 (4.30) 对应的增广矩阵为

$$\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 5 & 10 \end{array}\right].$$

应用例 1 中将矩阵 (4.25) 变换成矩阵 (4.28) 的行运算, 我们得到增广矩阵

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 5 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 0 & 31 \end{array}\right]. \tag{4.31}$$

当我们将最后一行表示为方程形式时将得到 $0 = 31$, 这是不可能的. 由于方程组 (4.30) 和 (4.31) 等价, 所以原方程组无解. □

在上述所有例子中, 方程的个数都不大于未知数的个数. 对方程个数比未知数个数大的方程组, 仍可用 Gauss-Jordan 消元法求解. 例如, 若我们考虑例 1 中的方

程组, 它的解为 $x = 124, y = 75, z = 31$. 如果我们将这组解满足的其他方程 (例如方程 $2x - 3y + z = 54$) 添加到该方程组中, 那么应用 Gauss-Jordan 消元法后我们将得到增广矩阵

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

它的最后一行全为零. 然而如果我们将这组解不满足的方程 (例如 $x + y + z = 1$) 添加到方程组中, 那么应用 Gauss-Jordan 消元法后我们将得到形如

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 124 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 31 \\ 0 & 0 & 0 & a \end{array}\right]$$

的增广矩阵, 其中 $a \neq 0$. 此时矩阵的最后一行给出一个矛盾的方程 $0 = a$, 这说明方程组无解.

4.19 方阵的逆

定义 (非奇异矩阵) *令 $\boldsymbol{A} = (a_{ij})$ 为 $n \times n$ 方阵, 如果有 $n \times n$ 矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{BA} = \boldsymbol{I}$, 其中 $\boldsymbol{I}$ 为 $n \times n$ 单位矩阵, 那么称 $\boldsymbol{A}$ 为***非奇异矩阵** (nonsingular matrix), *且称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的左逆.*

定理 4.20 (a) *如果 $n \times n$ 矩阵 $\boldsymbol{A}$ 有左逆 $\boldsymbol{B}$, 那么 $\boldsymbol{B}$ 也是 $\boldsymbol{A}$ 的右逆, 即 $\boldsymbol{AB} = \boldsymbol{I}$, 且 $\boldsymbol{A}$ 仅有一个左逆.*

(b) *如果 $n \times n$ 矩阵 $\boldsymbol{A}$ 有右逆 $\boldsymbol{B}$, 那么 $\boldsymbol{B}$ 也是 $\boldsymbol{A}$ 的左逆, 所以 $\boldsymbol{A}$ 非奇异.*

(a) 的证明 选单位坐标向量作为 $\mathbb{R}^n$ 的基, 并令 $T : \mathbb{R}^n \to \mathbb{R}^n$ 是以 $m(T) = \boldsymbol{A}$ 为矩阵表示的线性变换. 首先我们证明由 $T(x) = \boldsymbol{O}$ 可得 $\boldsymbol{x} = \boldsymbol{O}$, 从而证明 T 可逆.

任取一个使 $T(\boldsymbol{x}) = \boldsymbol{O}$ 的 x, 令 $\boldsymbol{X}$ 为由 $\boldsymbol{x}$ 的各分量组成的 $n \times 1$ 列矩阵. 由于 $T(\boldsymbol{x}) = \boldsymbol{O}$, 所以矩阵乘积 $\boldsymbol{AX}$ 是一个元素全为零的 $n \times 1$ 列矩阵, 所以对任意 $n \times n$ 矩阵 $\boldsymbol{B}$, $\boldsymbol{B}(\boldsymbol{AX})$ 也是一个元素全为零的 $n \times 1$ 列矩阵, 所以如果 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的左逆, 那么我们有

$$\boldsymbol{B}(\boldsymbol{AX}) = (\boldsymbol{BA})\boldsymbol{X} = \boldsymbol{IX} = \boldsymbol{X},$$

所以 $\boldsymbol{X}$ 的所有分量都为 0, 因此 T 可逆, 它的左逆唯一且同时也是 T 的右逆. 等式 $TT^{-1} = \boldsymbol{I}$ 表明 $m(T)m(T^{-1}) = \boldsymbol{I}$, 即 $\boldsymbol{A}m(T^{-1}) = \boldsymbol{I}$, 用 $\boldsymbol{B}$ 左乘此式两边可得 $m(T^{-1}) = \boldsymbol{B}$, 所以 $m(T)m(T^{-1}) = \boldsymbol{I}$ 表明 $\boldsymbol{AB} = \boldsymbol{I}$, 于是 $\boldsymbol{B}$ 也是 $\boldsymbol{A}$ 的右逆. 最后,

如果 $\boldsymbol{C}$ 也是 $\boldsymbol{A}$ 的左逆, 那么 $\boldsymbol{BA}=\boldsymbol{CA}=\boldsymbol{I}$, 用 $\boldsymbol{B}$ 右乘等式 $\boldsymbol{CA}=\boldsymbol{I}$ 的两边可知 $\boldsymbol{C}=\boldsymbol{B}$, 所以 $\boldsymbol{A}$ 仅有一个左逆.

(b) 的证明 如果 $\boldsymbol{AB}=\boldsymbol{I}$, 那么 $\boldsymbol{B}$ 有左逆 $\boldsymbol{A}$, 因此我们可对 $\boldsymbol{B}$ 用 (a) 中结论并推导得到 $\boldsymbol{A}$ 也是 $\boldsymbol{B}$ 的右逆, 所以 $\boldsymbol{BA}=\boldsymbol{I}$. 这表明 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的左逆, 所以 $\boldsymbol{A}$ 非奇异. □

定理 4.20 表明, 若一个方阵有左逆或右逆, 那么它是非奇异矩阵, 且此时它的左逆等于它的右逆 (可将矩阵与函数作比较, 函数可能有多个右逆, 也可能没有左逆). 如果 $\boldsymbol{A}$ 非奇异且 $\boldsymbol{B}$ 是它的左 (右) 逆, 那么 $\boldsymbol{BA}=\boldsymbol{AB}=\boldsymbol{I}$, 我们称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的逆并记作 $\boldsymbol{A}^{-1}$. 由于 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$, 所以 $\boldsymbol{A}^{-1}$ 也非奇异且 $\boldsymbol{A}$ 是它的逆. 另外我们也注意到, 当且仅当线性变换 T 的矩阵表示 $m(T)$ 可逆时, T 可逆.

如果非奇异 $n\times n$ 矩阵 $\boldsymbol{A}$ 的逆已知, 我们可用 $\boldsymbol{A}$ 的逆求解以 $\boldsymbol{A}$ 为系数矩阵的由 n 个 n 元方程组成的线性方程组. 例如, 方程组

$$\sum_{k=1}^{n} a_{ik}x_k = c_i, \quad i=1,2,\cdots,n,$$

可简单地写为矩阵方程

$$\boldsymbol{AX}=\boldsymbol{C},$$

其中 $\boldsymbol{A}=(a_{ij})$ 为系数矩阵, $\boldsymbol{X}$ 和 $\boldsymbol{C}$ 为下列矩阵:

$$\boldsymbol{X}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{C}=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

如果 $\boldsymbol{A}$ 非奇异, 那么方程有唯一解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{C}$. 这可看作初等代数中类似结果的推广, 在初等代数中线性方程 $ax=c$ 在 $a\neq 0$ 时有唯一解 $x=a^{-1}c$.

用公式 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{C}$ 解方程组看上去很简单, 其实却不然. 仅当 $\boldsymbol{A}^{-1}$ 已知时, 它才有用. 求 $\boldsymbol{A}^{-1}$ 的问题当 n 较大时可能令人难以招架. 事实上, 它与解 n 个独立的非齐次线性方程组等价. 设 $\boldsymbol{A}=(a_{ij})$ 非奇异, 令 $\boldsymbol{A}^{-1}=(b_{ij})$ 为它的逆, 如下 n^2 个方程给出了 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{-1}$ 的元之间的关系:

$$\sum_{k=1}^{n} a_{ik}b_{kj} = c_{ij}, \quad i,j=1,2,\cdots,n, \tag{4.32}$$

其中 c_{ij} 为单位矩阵的 ij 元: 若 $i=j$, 则 $c_{ij}=1$; 若 $i\neq j$, 则 $c_{ij}=0$. 对每一个 j, 我们都可将 (4.32) 看作一个包含 n 个线性方程和 n 个未知数 $b_{1j},b_{2j},\cdots,b_{nj}$ 且系数矩阵为 $\boldsymbol{A}$ 的非齐次方程组. 由于 $\boldsymbol{A}$ 非奇异, 每一个这样的方程组都有唯一解, 即 $\boldsymbol{B}$ 的第 j 列. 所有这样的方程组的系数矩阵都为 $\boldsymbol{A}$, 它们的差别仅在等号右边. 例如, 若 $\boldsymbol{A}$ 是一个 3×3 矩阵, 那么此时 (4.32) 包含 9 个方程, 我们可将这 9 个方

程表示为增广矩阵是如下矩阵的 3 个独立的方程组:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 1 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \end{array}\right], \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 1 \\ a_{31} & a_{32} & a_{33} & 0 \end{array}\right], \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 1 \end{array}\right].$$

对它们应用 Gauss-Jordan 消元法, 我们将分别得到如下增广矩阵:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{11} \\ 0 & 1 & 0 & b_{21} \\ 0 & 0 & 1 & b_{31} \end{array}\right], \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{12} \\ 0 & 1 & 0 & b_{22} \\ 0 & 0 & 1 & b_{32} \end{array}\right], \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & b_{13} \\ 0 & 1 & 0 & b_{23} \\ 0 & 0 & 1 & b_{33} \end{array}\right].$$

在实际应用中, 我们将充分利用三个方程组的系数矩阵相同这一特点, 从而通过对合并后的增广矩阵

$$\left[\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}\right],$$

应用 Gauss-Jordan 消元法同时解这三个方程组. 我们也用更简洁的形式 $[\boldsymbol{A}|\boldsymbol{I}]$ 表示上式. Gauss-Jordan 消元法最终将给出增广矩阵

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 & b_{31} & b_{32} & b_{33} \end{array}\right],$$

我们将上式记作 $[\boldsymbol{I}|\boldsymbol{B}]$. 竖线左边的矩阵 3×3 单位矩阵, 右边的矩阵为 $\boldsymbol{A}$ 的逆 (下一个定理将证明此结论).

顺便提及, 应用 Gauss-Jordan 消元法并不要求我们事先明确 $\boldsymbol{A}$ 非奇异. 如果 $\boldsymbol{A}$ **奇异**(singular), 即不是非奇异, 我们仍然可以应用 Gauss-Jordan 消元法, 只是在计算的过程中, 我们将会看到竖线右边的矩阵的对角线上的某个元素为零, 所以不可能将 $\boldsymbol{A}$ 转化为单位矩阵.

下面的定理解释了为什么应用 Gauss-Jordan 消元法可得到非奇异矩阵的逆.

定理 4.21 *令 $\boldsymbol{A}$ 为一个 $n \times n$ 矩阵且 $\boldsymbol{I}$ 为 $n \times n$ 单位矩阵, 如果 Gauss-Jordan 消元法将增广矩阵*

$$[\boldsymbol{A}|\boldsymbol{I}] \tag{4.33}$$

化为形如

$$[\boldsymbol{I}|\boldsymbol{B}] \tag{4.34}$$

的增广矩阵, 那么 $\boldsymbol{A}$ 非奇异且 $\boldsymbol{B} = \boldsymbol{A}^{-1}$.

证明 在 Gauss-Jordan 消元法中, 我们只用到三种初等行运算:

(1) 交换两行;

(2) 用一个非零纯量乘某一行;

(3) 将某一行的纯量倍数加到另一行上.

我们定义**初等矩阵** (elementary matrix) 为对单位矩阵应用任意一个初等行变换后得到的矩阵. 例如, 令 $\boldsymbol{E}$ 表示交换单位矩阵的前两行得到的初等矩阵. 当考虑 3×3 矩阵时,

$$\boldsymbol{E}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}.$$

取任意 3×3 矩阵 $\boldsymbol{A}=(a_{ij})$ 并用 $\boldsymbol{E}$ 左乘 $\boldsymbol{A}$, 我们得到

$$\boldsymbol{EA}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{21}&a_{22}&a_{23}\\a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\end{bmatrix}.$$

也就是说, 它们的积 $\boldsymbol{EA}$ 等于交换 $\boldsymbol{A}$ 的前两行所得的矩阵. 这条性质对所有矩阵都成立, 如果用交换单位矩阵的两行所得的初等矩阵 $\boldsymbol{E}$ 左乘任意矩阵 $\boldsymbol{A}$, 那么它们的积 $\boldsymbol{EA}$ 等于交换 $\boldsymbol{A}$ 中相应的两行所得的矩阵.

对其他两种初等行运算我们也有相似的结论. 例如, 若用非零纯量 c 乘 3×3 单位矩阵的第二行, 那么我们将得到矩阵

$$\boldsymbol{E}=\begin{bmatrix}1&0&0\\0&c&0\\0&0&1\end{bmatrix}.$$

若用该初等矩阵 $\boldsymbol{E}$ 左乘任意 3×3 矩阵 $\boldsymbol{A}=(a_{ij})$, 那么我们有

$$\boldsymbol{EA}=\begin{bmatrix}1&0&0\\0&c&0\\0&0&1\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\ca_{21}&ca_{22}&ca_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix},$$

它等于用 c 乘 $\boldsymbol{A}$ 的第二行所得的矩阵.

类似地, 当我们将用 c 乘 3×3 单位矩阵的第三行得到的结果加到第一行时, 会得到初等矩阵

$$\boldsymbol{E}=\begin{bmatrix}1&0&c\\0&1&0\\0&0&1\end{bmatrix}.$$

若用该初等矩阵 $\boldsymbol{E}$ 左乘任意 3×3 矩阵 $\boldsymbol{A}=(a_{ij})$, 我们有

$$\boldsymbol{EA}=\begin{bmatrix}1&0&c\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}a_{11}+ca_{31}&a_{12}+ca_{32}&a_{13}+ca_{33}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix},$$

它等于将用 c 乘 $\boldsymbol{A}$ 的第三行得到的结果加到第一行后所得的矩阵.

这些例子说明了一条普遍而又容易验证的性质：

如果用初等矩阵 $\boldsymbol{E}$ 左乘方阵 $\boldsymbol{A}$, 那么它们的积 $\boldsymbol{EA}$ 等于对 $\boldsymbol{A}$ 进行与将 $\boldsymbol{I}$ 化为 $\boldsymbol{E}$ 的初等行变换相同的变换后所得的矩阵.

易证该性质对任意 $\boldsymbol{E}$ 和 $\boldsymbol{A}$ 都成立.

当我们对增广矩阵 $[\boldsymbol{A}|\boldsymbol{I}]$ 作一系列初等行变换时, 就得到一系列形如

$$[\boldsymbol{E}_1\boldsymbol{A}|\boldsymbol{E}_1\boldsymbol{I}],\quad [\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{A}|\boldsymbol{E}_2\boldsymbol{E}_1\boldsymbol{I}],\quad ,\cdots,\quad [\boldsymbol{E}_k\cdots\boldsymbol{E}_1\boldsymbol{A}|\boldsymbol{E}_k\cdots\boldsymbol{E}_1\boldsymbol{I}],$$

的增广矩阵, 其中 $\boldsymbol{E}_1,\cdots,\boldsymbol{E}_k$ 为与所用的初等行变换对应的初等矩阵. 如果乘积 $\boldsymbol{E}_k\cdots\boldsymbol{E}_1\boldsymbol{A}$ 等于单位矩阵 $\boldsymbol{I}$, 那么说明 $\boldsymbol{A}$ 非奇异, 且乘积 $\boldsymbol{E}_k\cdots\boldsymbol{E}_1$ 等于 $\boldsymbol{A}^{-1}$. □

4.20　习　　题

对下列各方程组应用 Gauss-Jordan 消元法, 如果方程组有解, 那么求它的通解.

1. $\begin{cases} x+y+3z=5 \\ 2x-y+4z=11 \\ -y+z=3. \end{cases}$

2. $\begin{cases} 3x+2y+z=1 \\ 5x+3y+3z=2 \\ x+y-z=1. \end{cases}$

3. $\begin{cases} 3x+2y+z=1 \\ 5x+3y+3z=2 \\ 7x+4y+5z=3. \end{cases}$

4. $\begin{cases} 3x+2y+z=1 \\ 5x+3y+3z=2 \\ 7x+4y+5z=3 \\ x+y-z=0. \end{cases}$

5. $\begin{cases} 3x-2y+5z+u=1 \\ x+y-3z+2u=2 \\ 6x+y-4z+3u=7. \end{cases}$

6. $\begin{cases} x+y-3z+u=5 \\ 2x-y+z-2u=2 \\ 7x+y-7z+3u=3. \end{cases}$

7. $\begin{cases} x+y+2z+3u+4v=0 \\ 2x+2y+7z+11u+14v=0 \\ 3x+3y+6z+10u+15v=0. \end{cases}$

8. $\begin{cases} x-2y+z+2u=-2 \\ 2x+3y-z-5u=9 \\ 4x-y+z-u=5 \\ 5x-3y+2z+u=3. \end{cases}$

9. 证明当 $a\neq 8$ 时, 方程组 $x+y+2z=2,\ 2x-y+3z=2,\ 5x-y+az=6$ 有唯一解. 求该方程组在 $a=8$ 时的所有解.

10. (a) 求方程组

$$\begin{cases} 5x+2y-6z+2u=-1 \\ x-y+z-u=-2. \end{cases}$$

的所有解.

(b) 求方程组

$$\begin{cases} 5x+2y-6z+2u=-1 \\ x-y+z-u=-2. \\ x+y+z=6. \end{cases}$$

的所有解.

11. 本题表明如何求出所有 2×2 非奇异矩阵. 证明

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}=(ad-bc)I.$$

并由此证明命题：当且仅当 $ad-bc\neq 0$ 时，$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 非奇异，此时它的逆为

$$\frac{1}{(ad-bc)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

求习题 12～16 中矩阵的逆.

12. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}$.

13. $\begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix}$.

14. $\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{bmatrix}$.

15. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

16. $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{bmatrix}$.

4.21 关于矩阵的综合性习题

1. 若一个方阵包含一个全零行或一个全零列，证明它奇异.
2. 证明或举反例推翻下列各个关于 $n\times n$ 矩阵的命题.
   (a) 若 $\boldsymbol{AB}+\boldsymbol{BA}=\boldsymbol{O}$，则 $\boldsymbol{A}^2\boldsymbol{B}^3=\boldsymbol{B}^3\boldsymbol{A}^2$.
   (b) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 非奇异，则 $\boldsymbol{A}+\boldsymbol{B}$ 非奇异.
   (c) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 非奇异，则 $\boldsymbol{AB}$ 非奇异.
   (d) 若乘积 $\boldsymbol{AB}$ 非奇异，则 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都非奇异.
   (e) 若 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{A}+\boldsymbol{B}$ 都非奇异，则 $\boldsymbol{A}-\boldsymbol{B}$ 非奇异.
   (f) 若 $\boldsymbol{A}^2=\boldsymbol{O}$，则 $\boldsymbol{A}=\boldsymbol{O}$.
   (g) 若 $\boldsymbol{A}^2=\boldsymbol{O}$，则 $\boldsymbol{A}+\boldsymbol{I}$ 非奇异.
   (h) 若 $\boldsymbol{A}^3=\boldsymbol{O}$，则 $\boldsymbol{A}-\boldsymbol{I}$ 非奇异.
   (i) 若 $(\boldsymbol{A}+2\boldsymbol{I})^2=\boldsymbol{O}$，则 $\boldsymbol{A}$ 非奇异.
   (j) 若 $\boldsymbol{AB}=\boldsymbol{BA}$，则对所有整数 $k\geqslant 1$，都有 $\boldsymbol{AB}^k=\boldsymbol{B}^k\boldsymbol{A}$.
3. 若 $\boldsymbol{A}=\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$，求非奇异矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{bmatrix} 6 & 0 \\ 0 & -1 \end{bmatrix}$.

   提示：考虑使 $\boldsymbol{AP}=\boldsymbol{P}\begin{bmatrix} 6 & 0 \\ 0 & -1 \end{bmatrix}$ 的元素待定的矩阵 $\boldsymbol{P}$.
4. 已知矩阵 $\boldsymbol{A}=\begin{bmatrix} a & i \\ i & b \end{bmatrix}$ 具有性质 $\boldsymbol{A}^2=\boldsymbol{A}$，其中 $i^2=-1$，$a=\frac{1}{2}(1+\sqrt{5})$，$b=\frac{1}{2}(1-\sqrt{5})$，求具有性质 $\boldsymbol{A}^2=\boldsymbol{A}$ 且元为复数的所有 2×2 矩阵 $\boldsymbol{A}$.
5. 若 $\boldsymbol{A}^2=\boldsymbol{A}$，证明 $(\boldsymbol{A}+\boldsymbol{I})^k=\boldsymbol{I}+(2^k-1)\boldsymbol{A}$.
6. 狭义相对论用到形如 $x'=a(x-vt)$，$y'=y$，$z'=z$，$t'=a(t-vx/c^2)$ 的微分方程组，其中 v 代表某运动物体的速度，c 代表光速，且 $a=c/\sqrt{c^2-v^2}$，其中 $|v|<c$. 为了纪念荷兰物理学家 Hendrik A. Lorentz(1853—1928)，把将二维向量 (x,t) 映为 (x',t') 的线性变换称为 **Lorentz 变换**，我们将

该变换关于单位坐标向量的矩阵表示记为 $L(v)$,

$$L(v) = a\begin{bmatrix} 1 & -v \\ -vc^{-2} & 1 \end{bmatrix},$$

注意到 $L(v)$ 非奇异且 $L(0) = \boldsymbol{I}$, 证明 $L(v)L(u) = L(w)$, 其中 $w = (u+v)c^2/(uv+c^2)$. 也就是说, 两个 Lorentz 变换之积是一个 Lorentz 变换.

7. 将矩阵 $\boldsymbol{A}$ 的行和列互换得到的新矩阵称为 $\boldsymbol{A}$ 的**转置** (transpose), 记作 $\boldsymbol{A}^{\mathrm{T}}$ (读作 $\boldsymbol{A}$ 转置). 因此如果

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad 那么 \quad \boldsymbol{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

证明转置具有如下性质.

(a) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$.

(b) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$.

(c) $(c\boldsymbol{A})^{\mathrm{T}} = c\boldsymbol{A}^{\mathrm{T}}$.

(d) $(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$. (注意矩阵顺序的变化)

(e) $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$, 若 $\boldsymbol{A}$ 非奇异.

8. 当方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{AA}^{\mathrm{T}} = \boldsymbol{I}$ 时, 称 $\boldsymbol{A}$ 为**正交矩阵**.

(a) 证明对任意实数 θ, 2×2 矩阵 $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ 都是正交矩阵.

(b) 若 $\boldsymbol{A}$ 是任意实 $n\times n$ 矩阵, 证明当我们将它的行看作是 $\mathbb{R}^n$ 中的向量时, 这些行组成一个标准正交组.

9. 证明或举反例推翻下列各题中关于 $n\times n$ 矩阵的命题.

(a) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是正交矩阵, 则 $\boldsymbol{A}+\boldsymbol{B}$ 也是正交矩阵.

(b) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是正交矩阵, 则 $\boldsymbol{AB}$ 也是正交矩阵.

(c) 若 $\boldsymbol{A}$ 和 $\boldsymbol{AB}$ 都是正交矩阵, 则 $\boldsymbol{B}$ 也是正交矩阵.

10. 以法国数学家 Jacques Hadamard (1865—1963) 的名字命名的矩阵——**Hadamard 矩阵**是具有如下三条性质的 $n\times n$ 矩阵:

I. 所有元都只能取 1 或 -1;

II. 当我们将矩阵的行看作是 $\mathbb{R}^n$ 中的向量时, 各行的长度都为 $\sqrt{n}$;

III. 任意两个相异行的点积为 0.

几何与数论的一些问题将会用到 Hadamard 矩阵, 它在用于空间通讯的最优码中也有应用.

尽管 Hadamard 矩阵形式简单, 然而它尚有很多未解决的问题. 现在最重要的一个有待解决的问题就是求使得 $n\times n$ Hadamard 矩阵存在的所有 n. 本题给出这个问题的一部分解.

(a) 求所有 2×2 Hadamard 矩阵 (共有 8 个).

(b) 这部分给出如下定理的一个简单证明:

若 $\boldsymbol{A}$ 是 $n\times n$ Hadamard 矩阵, 其中 $n>2$, 那么 n 是 4 的倍数.

本证明基于两条关于 $\mathbb{R}^n$ 中向量的简单引理. 分别证明这两条引理并将它们用于 Hadamard 矩阵的行从而证明上述定理.

引理 4.22　*若 $\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}$ 为 $\mathbb{R}^n$ 中的正交向量, 那么我们有*

$$(\boldsymbol{X}+\boldsymbol{Y})\cdot(\boldsymbol{X}+\boldsymbol{Z}) = \|\boldsymbol{X}\|^2. \qquad \square$$

引理 4.23　*令 $\boldsymbol{X} = (x_1,\cdots,x_n)$, $\boldsymbol{Y} = (y_1,\cdots,y_n)$, $\boldsymbol{Z} = (z_1,\cdots,z_n)$, 若所有分量 x_i, y_i, z_i 都为 1 或 -1, 那么任意乘积 $(x_i+y_i)(x_i+z_i)$ 要么等于 0 要么等于 4.* □

第5章 行 列 式

5.1 引 言

在线性代数对于几何与分析的应用中，行列式的概念起着重要作用. 本章研究行列式的基本性质及其应用.

二阶和三阶行列式已经在第 2 章中作为将某些公式表示为紧凑形式的合适记号而引入. 具体言之，二阶行列式由公式

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{5.1}$$

所定义.

尽管表示法十分相似，但是行列式 $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ 与矩阵 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 是两个完全不同的概念. 行列式是依照公式 (5.1) 指定给矩阵的一个数. 为了强调行列式与矩阵的这种关系，我们也常常写作

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

在第 2 章中，三阶行列式是通过下述公式用二阶行列式来定义的：

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{5.2}$$

本章讨论一般情形，即对任意正整数 $n \geqslant 1$，讨论 n 阶方阵的行列式. 对以 a_{11} 为元素的矩阵，我们定义它的行列式就是这个元素，即 $\det[a_{11}] = a_{11}$.

我们的观点是把行列式看作一个函数，亦即对每个方阵 $\boldsymbol{A}$ 指定一个数，这个数叫作 $\boldsymbol{A}$ 的行列式并记作 $\det \boldsymbol{A}$. 定义这个函数的一种方法是将公式 (5.1) 与公式 (5.2) 推广，以给出一个明确的表达式. 这个表达式包含了 $\boldsymbol{A}$ 中元素的 $n!$ 个乘积. 对于较大的 n，这个表达式十分复杂因而在实际上很少使用. 我们更愿意用另一种方法来研究行列式，这种方法强调行列式的本质属性. 这些属性在应用中十分重要，因此我们把它们当作行列式函数的**公理** (axiom).

我们的方案由三部分组成. (1) 说明选择这些公理的动机. (2) 从这些公理推出行列式的进一步性质. (3) 证明对每一个 n 都有且仅有一个函数满足这些公理.

5.2 行列式函数公理的选择

我们已经在第 2 章中证明, 在三维空间中, 3 个向量 $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$ 的纯量三重积可以表示为以这 3 个向量作为行向量的矩阵的行列式, 即

$$\boldsymbol{A}_1 \times \boldsymbol{A}_2 \cdot \boldsymbol{A}_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

其中 $\boldsymbol{A}_1 = (a_{11}, a_{12}, a_{13}), \boldsymbol{A}_2 = (a_{21}, a_{22}, a_{23})$ 以及 $\boldsymbol{A}_3 = (a_{31}, a_{32}, a_{33})$.

若这 3 个行向量线性无关, 则纯量三重积不为零, 并且其绝对值等于由向量 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 与 $\boldsymbol{A}_3$ 所确定的**平行六面体** (Parallelepiped) 的体积. (见图 5.1 所示之例.) 若行向量线性相关, 则纯量三重积等于零, 这时向量 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 与 $\boldsymbol{A}_3$ 共面, 因而它们所确定的平行六面体退化成一平面图形, 从而其体积为零.

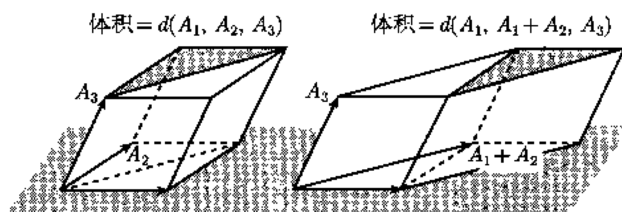


图 5.1　性质 $d(\boldsymbol{A}_1, \boldsymbol{A}_2 + \boldsymbol{A}_1, \boldsymbol{A}_3) = d(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$ 的几何解释. 这两个平行六面体的体积相同

纯量三重积的某些性质启发我们在高维情形如何选择关于行列式函数的公理. 为了用一种适于推广的形式来表述这些性质, 我们把行向量 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 与 $\boldsymbol{A}_3$ 的纯量三重积看作 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 与 $\boldsymbol{A}_3$ 的一个函数并记作 d, 即

$$d(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3) = \boldsymbol{A}_1 \times \boldsymbol{A}_2 \cdot \boldsymbol{A}_3.$$

我们特别关注以下性质.

(a) 若用纯量 t 去乘某一行, 则所得的纯量三重积是原来的 t 倍.

(b) 若将某一行用这一行与另一行的和代替, 则纯量三重积不变.

(c) 3 个单位坐标向量的纯量三重积 $\boldsymbol{i} \times \boldsymbol{j} \cdot \boldsymbol{k}$ 等于 1.

所有这些性质可以很容易地从点积和叉积导出. 例如, 若用 t 乘第一行, 则得 $d(t\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3) = td(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3)$ 对所有纯量 t 都成立. 若将第二行用这一行本身与另一行的和代替, 则性质 (b) 是说

$$d(\boldsymbol{A}_1, \boldsymbol{A}_2 + \boldsymbol{A}_1, \boldsymbol{A}_3) = d(\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3). \tag{5.3}$$

这个性质的几何解释见图 5.1 所示, 图 5.1 中分别给出了由 $\boldsymbol{A}_1, \boldsymbol{A}_2$ 与 $\boldsymbol{A}_3$ 确定的

平行六面体和由 $\boldsymbol{A}_1, \boldsymbol{A}_2+\boldsymbol{A}_1$ 与 $\boldsymbol{A}_3$ 确定的平行六面体. 等式 (5.3) 是说这两个平行六面体的体积相等. 从几何上看这一点很显然, 因为这两个平行六面体有相同的高和相同的底面积.

5.3 行列式函数的公理

上一节中所提到的纯量三重积的性质很容易推广并且可用作 n 阶行列式的公理. 若 $\boldsymbol{A}=(a_{ij})$ 是 $n\times n$ 实矩阵或复矩阵, 我们用 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n$ 表示它的各行. 于是 $\boldsymbol{A}$ 的第 i 行是 n 维线性空间中的一个向量:

$$\boldsymbol{A}_i=(a_{i1},a_{i2},\cdots,a_{in}).$$

我们将 $\boldsymbol{A}$ 的行列式看作 $\boldsymbol{A}$ 的一个函数, 或看作 $\boldsymbol{A}$ 的 n 个行向量 $\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n$ 的一个函数, 并将这个函数的值记作 $\det\boldsymbol{A}$ 或 $d(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$.

行列式函数的公理化定义 设 det 是 $n\times n$ 矩阵 $\boldsymbol{A}$ 的一个纯量数值函数, 如果它满足下面公理. 则称 det 是一个 n **阶的行列式函数**.

公理 1 (行齐次性) *若 $\boldsymbol{B}$ 是将矩阵 $\boldsymbol{A}$ 的某一行乘以一个纯量 t 所得的矩阵, 则 $\det\boldsymbol{B}=t\det\boldsymbol{A}$.* □

公理 2 (行相加不变性) *若 $\boldsymbol{B}$ 是将矩阵 $\boldsymbol{A}$ 中的某行加到另一行之后所得的矩阵, 则 $\det\boldsymbol{B}=\det\boldsymbol{A}$.* □

公理 3 (单位矩阵的行列式等于 1) *若用 $\boldsymbol{I}_k$ 表示单位坐标向量, 则*

$$d(\boldsymbol{I}_1,\cdots,\boldsymbol{I}_n)=1.$$

值得指出的是, 对给定的 n, 存在唯一适合这 3 条公理的函数 det. 这一结论的一个证明将在后面 (定理 5.6 与定理 5.20) 给出. 这里我们将首先从这些公理导出关于行列式的进一步性质.

定理 5.1 *若 $\boldsymbol{A}$ 的某一行是零向量, 则 $\det\boldsymbol{A}=0$*

证明 设第 k 行 $\boldsymbol{A}_k=\boldsymbol{O}$, 令 $\boldsymbol{B}$ 为从 $\boldsymbol{A}$ 将 $\boldsymbol{A}_k$ 乘以 -1 后所得的矩阵, 则 $\boldsymbol{B}=\boldsymbol{A}$, 从而 $\det\boldsymbol{B}=\det\boldsymbol{A}$. 另一方面, 由公理 1 我们得 $\det\boldsymbol{B}=-\det\boldsymbol{A}$, 所以 $\det\boldsymbol{A}=0$. □

定理 5.2 *若 $\boldsymbol{B}$ 是将矩阵 $\boldsymbol{A}$ 中某一行的纯量倍数加到另一行后所得的矩阵, 则 $\det\boldsymbol{B}=\det\boldsymbol{A}$*

证明 设 $j\neq k$, 若 $\boldsymbol{B}$ 是通过将行 $\boldsymbol{A}_k$ 的 t 倍加到行 $\boldsymbol{A}_j$ 上所得的矩阵, 即 $\boldsymbol{B}_j=\boldsymbol{A}_j+t\boldsymbol{A}_k$, 而对所有 $i\neq j$ 都有 $\boldsymbol{B}_i=\boldsymbol{A}_i$. 若 $t=0$, 则 $\boldsymbol{B}=\boldsymbol{A}$, 从而 $\det\boldsymbol{B}=\det\boldsymbol{A}$.

今设 $t\neq 0$, 令 $\boldsymbol{C}$ 为这样的一个矩阵, 它的第 k 行是 $t\boldsymbol{A}_k$, 而对所有 $i\neq k$, $\boldsymbol{A}$ 的第 i 行为 $\boldsymbol{A}_i$. 由公理 1 得 $\det\boldsymbol{C}=t\det\boldsymbol{A}$. 再令 D 为将 $\boldsymbol{C}$ 中的第 k 行, 即 $t\boldsymbol{A}_k$, 加到第 j 行 $\boldsymbol{A}_j$ 所得到的矩阵, 即 $D_k=t\boldsymbol{A}_k$, $D_j=\boldsymbol{A}_j+t\boldsymbol{A}_k$, 而对 $i\neq j$ 且 $i\neq k$,

都有 $D_i = \boldsymbol{A}_i$. 由公理 2 得 $\det D = \det \boldsymbol{C}$. 但如果我们将 D 的第 k 行乘以 $\frac{1}{t}$, 则得矩阵 $\boldsymbol{B}$, 从而由公理 1 得 $\det \boldsymbol{B} = (\frac{1}{t}) \det \boldsymbol{C} = \det \boldsymbol{A}$. □

定理 5.3 (a) *若 $\boldsymbol{B}$ 是交换 $\boldsymbol{A}$ 的两行 $\boldsymbol{A}_i$ 与 $\boldsymbol{A}_j$ 所得的矩阵, 此处 $i \neq j$, 则 $\det \boldsymbol{B} = -\det \boldsymbol{A}$.*

(b) *若矩阵的两行相等, 则行列式为零.*

证明 (a) 我们引入三个与 $\boldsymbol{A}$ 有相同行列式的矩阵 $\boldsymbol{C}, D$ 和 $\boldsymbol{E}$, 这里我们只写出这些矩阵的第 i 行和第 j 行:

$$\begin{aligned} &\boldsymbol{C}_i = \boldsymbol{A}_i + \boldsymbol{A}_j, \quad &&\boldsymbol{C}_j = \boldsymbol{A}_j, \\ &D_i = \boldsymbol{C}_i, &&D_j = \boldsymbol{C}_j + (-1)\boldsymbol{C}_i, \\ &\boldsymbol{E}_i = D_i + D_j, &&\boldsymbol{E}_j = D_j. \end{aligned}$$

$\boldsymbol{C}, D$ 与 $\boldsymbol{E}$ 的其余各行与 $\boldsymbol{A}$ 的相同. 由公理 2 我们有 $\det \boldsymbol{C} = \det \boldsymbol{A}$; 由定理 5.2 我们有 $\det D = \det \boldsymbol{C}$; 由公理 2 我们有 $\det \boldsymbol{E} = \det D$. 因此 $\boldsymbol{C}, D$ 与 $\boldsymbol{E}$ 的行列式都等于 $\det \boldsymbol{A}$. 将 D 与 $\boldsymbol{E}$ 的第 i 行与第 j 行用 $\boldsymbol{A}$ 的行表示, 我们得

$$D_i = \boldsymbol{A}_i + \boldsymbol{A}_j, \quad D_j = \boldsymbol{A}_j - (\boldsymbol{A}_i + \boldsymbol{A}_j) = -\boldsymbol{A}_i,$$

以及

$$\boldsymbol{E}_i = \boldsymbol{A}_j, \quad \boldsymbol{E}_j = -\boldsymbol{A}_i.$$

因此, 矩阵 $\boldsymbol{B}$ 可以从 $\boldsymbol{E}$ 将行 $\boldsymbol{E}_j$ 乘以 -1 而得, 于是得

$$\det \boldsymbol{B} = -\det \boldsymbol{E} = -\det \boldsymbol{A}.$$

从而证得 (a). 为证 (b), 交换相等的两行, 由 (a) 即得. □

我们指出, 在定理 5.1、定理 5.2 与定理 5.3 的证明中并未用到公理 3.

5.4 对角矩阵的行列式

现在我们已经有了足够多的信息来计算行列式, 当然这是在假定行列式函数存在的前提之下. 下面讨论一类特殊的矩阵 —— 对角矩阵.

定义 (对角矩阵) *如果方阵 (a_{ij}) 的所有非主对角线元素都等于零, 即当 $i \neq j$ 时都有 $a_{ij} = 0$, 则称之为对角矩阵.*

换言之, 下述形式的矩阵 $\boldsymbol{A}$ 叫对角矩阵:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

我们用记号 $\boldsymbol{A} = \operatorname{diag}(a_{11}, a_{22}, \cdots, a_{nn})$ 来表示这个对角矩阵. 下述定理告诉我们, 对角矩阵的行列式等于它的主对角线元素的乘积.

定理 5.4 若 $\boldsymbol{A} = \text{diag}\,(a_{11}, a_{22}, \cdots, a_{nn})$, 则

$$\det \boldsymbol{A} = a_{11}a_{22}\cdots a_{nn}. \tag{5.4}$$

证明 对角矩阵 $\boldsymbol{A}$ 可以依次用其主对角线元素 $a_{11}, a_{22}, \cdots, a_{nn}$ 去乘 $n \times n$ 单位阵 $\boldsymbol{I}$ 的各行而得. 重复应用公理 1 即得

$$\det \boldsymbol{A} = a_{11}a_{22}\cdots a_{nn} \det \boldsymbol{I}, \tag{5.5}$$

再由公理 3 即得结论. □

5.5 上三角形矩阵的行列式

解线性方程组的 Gauss-Jordan 消元法也是计算行列式的最好的方法之一, 这个方法是对矩阵作如下三类行初等变换:

(1) 交换两行;

(2) 用一个非零的纯量去乘某行的所有元素;

(3) 将某行的纯量倍数加到另一行.

由定理 5.3, 第一类行初等变换只改变这个矩阵的行列式的正负号; 由公理 1, 施行第二类行初等变换的结果是所得矩阵的行列式等于用这个非零纯量去乘原来矩阵的行列式; 由定理 5.2, 施行第三类行初等变换不改变行列式的值. 接下来我们要证明, 只用第三类行初等变换, 我们总可以把任意一个上三角形矩阵化为一个对角矩阵, 从而求出它的行列式.

定义 (上三角形矩阵) 形如

$$\boldsymbol{U} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix}, \tag{5.6}$$

即主对角线以下元素都是零的方阵叫**上三角形矩阵**. □

定理 5.5 上三角形矩阵的行列式等于它主对角线上各元素的乘积. 换言之, 若 $\boldsymbol{U}$ 是由式 (5.6) 所给出的上三角形矩阵, 则

$$\det \boldsymbol{U} = u_{11}u_{22}\cdots u_{nn} \tag{5.7}$$

证明 若 $u_{nn} = 0$, 则上三角形矩阵 $\boldsymbol{U}$ 的最后一行是零向量, 故由定理 5.1 得 $\det \boldsymbol{U} = 0$, 因此式 (5.7) 左右两边都等于零, 故结论成立. 假定 $u_{nn} \neq 0$, 我们用 Gauss-Jordan 消元法中的第三类初等变换, 将 $\boldsymbol{U}$ 的最后一行乘以 $-\frac{u_{1n}}{u_{nn}}$ 后加到第一行, 如此得到的新矩阵将 $\boldsymbol{U}$ 中的 u_{1n} 用 0 代替而保持其余元素不变. 反复施行第三类行初等变换, 可使最后一列中位于 u_{nn} 上方的元素都变成零而保持其余元素不变, 并且也保持行列式不变. 由此我们得到这样一个上三角形矩阵 $\boldsymbol{U}'$, 它的最后一列的前 $n-1$ 个元素都是零:

$$
U' = \begin{pmatrix} u_{11} & u_{12} & \cdots & 0 \\ 0 & u_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix},
$$

而且 $\det U' = \det U$. 如果 U' 主对角线上的倒数第二个元素 $u_{n-1,n-1} = 0$, 则 U' 的倒数第二个行向量是零向量, 因此 $\det U' = 0$, 此时等式 (5.7) 的两边都是零, 因而显然成立. 若 $u_{n-1,n-1} \neq 0$, 则可反复施行第三类行初等变换. 如果每一个主对角线元素 u_{kk} 都不等于零, 则最后我们得到一个对角矩阵, 它的行列式等于其主对角线元素的乘积. 若某个主对角线元素 $u_{kk} = 0$, 则可得到一个分量全是零的行向量, 因此其行列式为零, 等式 (5.7) 显然成立. □

5.6 用 Gauss-Jordan 消元法计算行列式

反复施行第三类行初等变换, 我们可以将任意一个方阵 A 变为一个上三角形矩阵 U, 而对于上三角形矩阵, 我们已经知道如何求它的行列式. 由于每施行一次第一类行初等变换仅改变行列式的正负号, 用非零纯量 c 施行一次第二类行初等变换所得的行列式等于原行列式与 c 的乘积, 施行第三类行初等变换则不改变行列式的值. 因此, 如果共作了 p 次第一类行初等变换, q 次第二类行初等变换并且所用的非零纯量为 $c_1, \cdots, c_q$, 则我们得 $\det U = (-1)^p c_1 \cdots c_q \det A$, 从而得

$$
\det A = (-1)^p (c_1 \cdots c_q)^{-1} \det U.
$$

我们指出, 仅用公理 1 与公理 2 就可导出上述公式. 公理 3 只是在由等式 (5.5) 导出公式 (5.4) 时才用到, 而上述等式证明中并不依赖于公理 3. 换言之, 仅仅利用公理 1 与公理 2 我们就已经证明了对任意 $n \times n$ 矩阵 A 都存在一个 (依赖于 A 的) 纯量 $c(A)$ 使得

$$
\det A = c(A) \det I. \tag{5.8}
$$

5.7 行列式函数的唯一性

利用由公理 1 和公理 2 导出的公式 (5.8), 我们可以证明至多存在一个 n 阶的行列式函数.

定理 5.6 (行列式的唯一性定理) 设 d 是满足关于 n 阶行列式函数的所有三条公理的一个函数, 如果 f 是另外一个满足公理 1 与公理 2 的函数, 则对任意 $n \times n$ 矩阵 A, 我们都有

$$
f(A) = d(A) f(I). \tag{5.9}
$$

从而, 若 f 也满足公理 3, 则 $f(A) = d(A)$.

证明 令 $g(\boldsymbol{A}) = f(\boldsymbol{A}) - d(\boldsymbol{A})f(\boldsymbol{I})$, 则 g 满足公理 1 与公理 2, 但不满足公理 3. 事实上, 由于 $d(\boldsymbol{I}) = 1$, 因此我们有 $g(\boldsymbol{I}) = f(\boldsymbol{I}) - d(\boldsymbol{I})f(\boldsymbol{I}) = 0$. 由于 g 满足公理 1 与公理 2, 于是我们可以将公式 (5.8) 用于 g, 这是因为公式 (5.8) 的证明仅与公理 1 及公理 2 有关. 因此存在纯量 $c(\boldsymbol{A})$ 使 $g(\boldsymbol{A}) = c(\boldsymbol{A})g(\boldsymbol{I})$. 从而由 $g(\boldsymbol{I}) = 0$ 可知 $g(\boldsymbol{A}) = 0$ 对所有 $\boldsymbol{A}$ 都成立, 于是得 $f(\boldsymbol{A}) = d(\boldsymbol{A})f(\boldsymbol{I})$. □

5.8 习　题

在下面各题中, 我们都假定了行列式函数的存在性, 三阶行列式可利用公式 (5.2) 来计算.

1. 计算下列各行列式.

(a) $\begin{vmatrix} 2 & 1 & 1 \\ 1 & 4 & -4 \\ 1 & 0 & 2 \end{vmatrix}$. (b) $\begin{vmatrix} 3 & 0 & 8 \\ 5 & 0 & 7 \\ -1 & 4 & 2 \end{vmatrix}$. (c) $\begin{vmatrix} a & 1 & 0 \\ 2 & a & 2 \\ 0 & 1 & a \end{vmatrix}$.

2. 设 $\det \begin{bmatrix} x & y & z \\ 3 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} = 1$, 计算下列各矩阵的行列式.

(a) $\begin{bmatrix} 2x & 2y & 2z \\ \frac{3}{2} & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. (b) $\begin{bmatrix} x & y & z \\ 3x+3 & 3y & 3z+2 \\ x+1 & y+1 & z+1 \end{bmatrix}$.

(c) $\begin{bmatrix} x-1 & y-1 & z-1 \\ 4 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}$.

3. (a) 证明 $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b)$.

(b) 对行列式 $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix}$ 与 $\begin{vmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}$ 给出相应的公式.

4. 设 $\boldsymbol{A}$ 为一个 $n \times n$ 矩阵. 证明对任意纯量 c 都有 $\det(c\boldsymbol{A}) = c^n \det(\boldsymbol{A})$

5. 将下列各矩阵化为上三角形矩阵, 然后求出它们的行列式.

(a) $\begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$. (b) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}$.

(c) $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix}$. (d) $\begin{bmatrix} a & 1 & 0 & 0 & 0 \\ 4 & a & 2 & 0 & 0 \\ 0 & 3 & a & 3 & 0 \\ 0 & 0 & 2 & a & 4 \\ 0 & 0 & 0 & 1 & a \end{bmatrix}$.

(e) $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}$.

6. 设 $\boldsymbol{U}$ 与 V 为两个 $n \times n$ 上三角形矩阵.
   (a) 证明：$\boldsymbol{U}+V$ 和 $\boldsymbol{U}V$ 都是上三角形矩阵.
   (b) 证明：$\det(\boldsymbol{U}V) = (\det \boldsymbol{U})(\det V)$
   (c) 证明：若 $\det \boldsymbol{U} \neq 0$, 则存在上三角形矩阵 $\boldsymbol{U}^{-1}$ 使 $\boldsymbol{U}\boldsymbol{U}^{-1} = \boldsymbol{I}$ 并有 $\det(\boldsymbol{U}^{-1}) = 1/\det \boldsymbol{U}$
7. 对下述上三角形矩阵 $\boldsymbol{A}$, 求 $\det \boldsymbol{A}, \det(\boldsymbol{A}^{-1})$ 和 $\boldsymbol{A}^{-1}$：

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

8. 设 $\boldsymbol{A} = (a_{ij})$ 为一个 $n \times n$ 方阵. 若 $\boldsymbol{A}$ 的主对角线以下的元素都等于零, 即当 $i < j$ 时都有 $a_{ij} = 0$, 则称 $\boldsymbol{A}$ 为一个**下三角形矩阵**. 证明：下三角形矩阵的行列式等于它的全体主对角线元素的乘积：$\det \boldsymbol{A} = a_{11}a_{22}\cdots a_{nn}$.

   习题 9~11 需要用到微积分知识.

9. 设 f_1, f_2, g_1, g_2 为区间 (a,b) 上的四个可微函数. 对任意 $x \in (a,b)$, 定义 $F(x)$ 如下：

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix},$$

   证明：

$$F'(x) = \begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & f_2(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}.$$

10. 对行列式

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix}$$

   推广习题 9 中的结论并给出证明.

11. (a) 设 $F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix}$, 证明：$F'(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1''(x) & f_2''(x) \end{vmatrix}$.
   (b) 对 3×3 行列式叙述并证明相应的结论, 可以应用公式 (5.2).

5.9　行列式的多重线性性

作为唯一性定理的一个应用, 我们将导出行列式的一个重要性质——**多重线性性**(Multilinearity). 我们从下述特殊情形开始.

定理 5.7 (行可加性)　给定一个 $n \times n$ 矩阵 $\boldsymbol{A} = (\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n)$ 和一个 n 维向量 $\boldsymbol{V}$, 令 $\boldsymbol{B}$ 为将 $\boldsymbol{A}$ 中某行 $\boldsymbol{A}_k$ 用 V 代替后所得的矩阵, $\boldsymbol{C}$ 为从 $\boldsymbol{A}$ 中将 $\boldsymbol{A}_k$ 用

$\boldsymbol{A}_k + V$ 代替后所得的矩阵, 则

$$\det \boldsymbol{C} = \det \boldsymbol{A} + \det \boldsymbol{B}. \tag{5.10}$$

这一性质叫作关于第 k 行的可加性, 它可表述为如下形式:

$$d(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k + V, \cdots, \boldsymbol{A}_n) = d(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k, \cdots, \boldsymbol{A}_n) + d(\boldsymbol{A}_1, \cdots, V, \cdots, \boldsymbol{A}_n)$$

证明 令 $f(\boldsymbol{A}) = \det \boldsymbol{C} - \det \boldsymbol{B}$, 我们要证明 f 满足关于行列式函数的所有三条公理, 然后由唯一性定理可知必有 $f(\boldsymbol{A}) = \det \boldsymbol{A}$, 由此即得公式 (5.10).

若将 $\det \boldsymbol{C}$ 与 $\det \boldsymbol{B}$ 看作 $\boldsymbol{A}$ 的 n 个行的函数, 则它们都满足公理 1 与公理 2, 从而它们的差 $f(\boldsymbol{A})$ 也满足这两条公理. 剩下还需验证 f 也满足公理 3: $f(\boldsymbol{I}) = 1$, 此处 $\boldsymbol{I}$ 为单位矩阵. 在 f 的定义中取 $\boldsymbol{A} = \boldsymbol{I}$, 则得

$$f(\boldsymbol{I}) = \det(\boldsymbol{I}_1, \cdots, \boldsymbol{I}_k + V, \cdots, \boldsymbol{I}_n) - \det(\boldsymbol{I}_1, \cdots, V, \cdots, \boldsymbol{I}_n). \tag{5.11}$$

将 V 表为一个 n 序组, 设为 $V = (v_1, \cdots, v_n)$. 若 $k = 1$, 则 $\boldsymbol{C}$ 与 $\boldsymbol{B}$ 都是上三角形矩阵且它们的行列式分别为 $1 + v_1$ 与 v_1, 因此

$$f(\boldsymbol{I}) = \det \boldsymbol{C} - \det \boldsymbol{B} = (1 + v_1) - v_1 = 1.$$

若 $k = 2$, 则矩阵 $\boldsymbol{C}$ 与 $\boldsymbol{B}$ 的第二行分别为

$$\boldsymbol{C}_2 = (v_1, 1 + v_2, v_3, \cdots, v_n), \quad \boldsymbol{B}_2 = (v_1, v_2, v_3, \cdots, v_n),$$

其余各行都与 $\boldsymbol{I}$ 的相同. 对矩阵 $\boldsymbol{C}$ 与 $\boldsymbol{B}$ 施行第三类行初等变换, 可将它们化为上三角形矩阵并且保持其行列式不变. 实际上, 如果我们分别将 $\boldsymbol{C}$ 与 $\boldsymbol{B}$ 的第一行乘以 $-v_1$ 之后加到第二行则新的第二行将分别变为

$$(0, 1 + v_2, v_3, \cdots, v_n), \quad (0, v_2, v_3, \cdots, v_n).$$

于是我们得到行列式分别为 $(1 + v_2)$ 与 v_2 的两个上三角形矩阵, 因此它们的行列式之差 $f(\boldsymbol{I})$ 等于 1.

在一般情形, 对 $k \geqslant 2$, 可作同样的讨论. 经过一系列第三类行初等变换, 我们可将等式 (5.11) 右边的两个矩阵化为上三角形矩阵, 并且其行列式分别为 $1 + v_k$ 与 v_k, 因此 $f(\boldsymbol{I}) = 1$ 从而等式 (5.10) 成立.

若我们不是用一个向量 V, 而是用两个向量 $\boldsymbol{U}$ 与 V 的和 $V + \boldsymbol{U}$, 则应用定理 5.7 两次可得

$$\begin{aligned} d(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k + V + \boldsymbol{U}, \cdots, \boldsymbol{A}_n) &= d(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_k, \cdots, \boldsymbol{A}_n) \\ &+ d(\boldsymbol{A}_1, \cdots, V, \cdots, \boldsymbol{A}_n) + d(\boldsymbol{A}_1, \cdots, \boldsymbol{U}, \cdots, \boldsymbol{A}_n). \end{aligned}$$

一般来说, 我们可以同时用任意多个向量并将每个向量分别乘以一个纯量 (利用齐性), 由此得到一个性质, 称为对各行的线性性. 线性性将齐性与可加性结合在一起. 例如, 第一行的线性性可表示为:

$$d\left(\sum_{i=1}^{p} t_i \boldsymbol{U}_i, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n\right) = \sum_{i=1}^{p} t_i d(\boldsymbol{U}_i, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n),$$

此处 $t_1, \cdots, t_p$ 为任意纯量, $\boldsymbol{U}_1, \cdots, \boldsymbol{U}_p$ 为任意 n 维行向量. 由于这个性质对每

一行都成立, 因此我们常常说, 行列式是它的各行的一个**多重线性函数** (multilinear function). □

5.10 多重线性性的应用

我们的第一个应用是定理 5.3(b) 的下述推广.

定理 5.8 *若矩阵 $\boldsymbol{A}$ 的各行线性相关, 则 $d(\boldsymbol{A})=0$.*

证明 假定存在不全为零的纯量 $c_1,\cdots,c_n$ 使 $\sum\limits_{i=1}^{n}c_i\boldsymbol{A}_i=\boldsymbol{O}$, 则满足 $c_k\neq 0$ 的每一行都可表示为其余各行的线性组合. 为简便记, 不妨设第一行 $\boldsymbol{A}_1$ 可表示为其余各行的线性组合, 设为 $\boldsymbol{A}_1=\sum\limits_{i=2}^{n}t_i\boldsymbol{A}_i$. 由关于第一行的线性性, 我们有

$$d(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n)=d\left(\sum_{i=2}^{n}t_i\boldsymbol{A}_i,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n\right)=\sum_{i=2}^{n}t_id(\boldsymbol{A}_i,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n).$$

由于 $i\geqslant 2$, 因此 $\boldsymbol{A}_i$ 至少与 $\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 中的一行相等. 因此由定理 5.3(b) 可知上面这个和式中的每一项都等于零, 从而它们的和为零.

若 $\boldsymbol{A}_k$ 是其余各行的线性组合, 由关于第 k 行的线性性可得同样的结论. 因此我们已经证明: 若一个方阵的各行线性相关, 则它们的行列式等于零. □

作为多重线性性的第二个应用, 我们证明二阶与三阶方阵的行列式函数的存在性, 由此也给出了前面给出的公式 (5.1) 与 (5.2) 的合理性.

定理 5.9 (a) *对 2×2 矩阵, 存在唯一的行列式函数:*

$$\det\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}=\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}.$$

(b) *对 3×3 矩阵, 存在唯一的行列式函数:*

$$\det\begin{bmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}=a_{11}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix}+a_{13}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}. \tag{5.12}$$

证明 最简单的方法是验证定理中的两个公式都满足关于行列式函数的所有公理, 然后再由唯一性定理即得结论. 不过, 为了更好地理解这些公式的来源, 我们将用多重线性性直接推导出这两个公式.

我们从任意 2×2 矩阵 $\boldsymbol{A}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$ 出发并将它的行向量表示为单位坐标向量 $\boldsymbol{i}=(1,0)$ 与 $\boldsymbol{j}=(0,1)$ 的线性组合:

$$\boldsymbol{A}_1=(a_{11},a_{12})=a_{11}\boldsymbol{i}+a_{12}\boldsymbol{j},\quad \boldsymbol{A}_2=(a_{21},a_{22})=a_{21}\boldsymbol{i}+a_{22}\boldsymbol{j}$$

由关于第一行的线性性, 我们得

$$d(\boldsymbol{A}_1,\boldsymbol{A}_2)=d(a_{11}\boldsymbol{i}+a_{12}\boldsymbol{j},\boldsymbol{A}_2)=a_{11}d(\boldsymbol{i},\boldsymbol{A}_2)+a_{12}d(\boldsymbol{j},\boldsymbol{A}_2),$$

再由关于第二行的线性性, 因为 $d(\boldsymbol{i},\boldsymbol{i})$ 中的两行相同, 所以 $d(i,i)=0$, 从而得

$$d(\boldsymbol{i},\boldsymbol{A}_2)=d(\boldsymbol{i},a_{21}\boldsymbol{i}+a_{22}\boldsymbol{j})=a_{21}d(\boldsymbol{i},\boldsymbol{i})+a_{22}d(\boldsymbol{i},\boldsymbol{j})=a_{22}d(\boldsymbol{i},\boldsymbol{j}).$$

同理可得

$$d(\boldsymbol{j},\boldsymbol{A}_2)=d(\boldsymbol{j},a_{21}\boldsymbol{i}+a_{22}\boldsymbol{j})=a_{21}d(\boldsymbol{j},\boldsymbol{i})+a_{22}d(\boldsymbol{j},\boldsymbol{j})=-a_{21}d(\boldsymbol{i},\boldsymbol{j}),$$

其中正负号的改变是由于我们交换了 $d(\boldsymbol{j},\boldsymbol{i})$ 中的两行 $\boldsymbol{i}$ 与 $\boldsymbol{j}$. 因此得

$$d(\boldsymbol{A}_1,\boldsymbol{A}_2)=(a_{11}a_{22}-a_{12}a_{21})d(\boldsymbol{i},\boldsymbol{j}),$$

而由公理 3 得 $d(\boldsymbol{i},\boldsymbol{j})=1$, 因此得 $d(\boldsymbol{A}_1,\boldsymbol{A}_2)=(a_{11}a_{22}-a_{12}a_{21})$, 即得结论. □

同理可得关于 3 阶行列式的公式 (5.12), 我们把证明留作习题.

5.11 行列式的乘积公式

在本节中, 我们利用唯一性定理证明: 若行列式函数存在, 两个方阵乘积的行列式等于它们的行列式的乘积:

$$\det(\boldsymbol{AB})=(\det\boldsymbol{A})(\det\boldsymbol{B}).$$

由定义, 两个矩阵 $\boldsymbol{A}=(a_{ij})$ 与 $\boldsymbol{B}=(b_{ij})$ 的积是一个方阵 $\boldsymbol{C}=(c_{ij})$, 其中元素 c_{ij} 由公式

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj} \tag{5.13}$$

给出. 矩阵乘法仅当左边矩阵 $\boldsymbol{A}$ 的列数等于右边矩阵 $\boldsymbol{B}$ 的行数时才有意义, 当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个同阶方阵时这一条件自然满足.

乘积公式的证明需要用到 $\boldsymbol{AB}$ 的行与 $\boldsymbol{A}$ 的行之间的一个简单关系, 我们把它表述为下述引理, 如同往常一样, $\boldsymbol{A}_i$ 表示矩阵 $\boldsymbol{A}$ 的第 i 行.

引理 5.10 若 $\boldsymbol{A}$ 是一个 $m\times n$ 矩阵, $\boldsymbol{B}$ 是一个 $n\times p$ 矩阵, 则我们有

$$(\boldsymbol{AB})_i=\boldsymbol{A}_i\boldsymbol{B}, \tag{5.14}$$

即乘积 $\boldsymbol{AB}$ 的第 i 行等于矩阵 $\boldsymbol{A}$ 的第 i 行 $\boldsymbol{A}_i$ 与 $\boldsymbol{B}$ 的乘积.

证明 令 $\boldsymbol{B}^j$ 表示 $\boldsymbol{B}$ 的第 j 列, 并令 $\boldsymbol{C}=\boldsymbol{AB}$. 则等式 (5.13) 中右边的和可以看作 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 j 列的点积, 即

$$c_{ij}=\boldsymbol{A}_i\cdot\boldsymbol{B}^j.$$

因此 $\boldsymbol{C}$ 的第 i 行 $\boldsymbol{C}_i$ 是以各点积为元素的行矩阵:

$$\boldsymbol{C}_i=[\boldsymbol{A}_i\cdot\boldsymbol{B}^1,\boldsymbol{A}_i\cdot\boldsymbol{B}^2,\cdots,\boldsymbol{A}_i\cdot\boldsymbol{B}^p],$$

而这也正是行矩阵 $\boldsymbol{A}_i$ 与矩阵 $\boldsymbol{B}$ 的矩阵乘积:

$$\boldsymbol{A}_i\boldsymbol{B}=[a_{i1},\cdots,a_{in}]\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}=[\boldsymbol{A}_i\cdot\boldsymbol{B}^1,\boldsymbol{A}_i\cdot\boldsymbol{B}^2,\cdots,\boldsymbol{A}_i\cdot\boldsymbol{B}^p],$$

从而 $\boldsymbol{C}_i=\boldsymbol{A}_i\boldsymbol{B}$, 即得引理. □

定理 5.11(行列式乘积公式) *设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为两个 n 阶方阵, 则*

$$\det(\boldsymbol{AB})=(\det\boldsymbol{A})(\det\boldsymbol{B}). \tag{5.15}$$

证明 由引理 5.10, 我们有 $(\boldsymbol{AB})_i=\boldsymbol{A}_i\boldsymbol{B}$, 因此

$$\det(\boldsymbol{AB})=d(\boldsymbol{A}_1\boldsymbol{B},\cdots,\boldsymbol{A}_n\boldsymbol{B}).$$

固定 $\boldsymbol{B}$, 令 $f(\boldsymbol{A})=\det(\boldsymbol{AB})$ 并将它表示为 $\boldsymbol{A}$ 的各行的一个函数, 我们有

$$f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)=d(\boldsymbol{A}_1\boldsymbol{B},\cdots,\boldsymbol{A}_n\boldsymbol{B}).$$

容易验证 f 满足关于行列式函数的前两条公理. 因此, 由唯一性定理, 我们得 $f(\boldsymbol{A})=d(\boldsymbol{A})f(\boldsymbol{I})$. 在 f 的定义中取 $\boldsymbol{A}=\boldsymbol{I}$, 则得 $f(\boldsymbol{I})=d(\boldsymbol{B})$, 因此 $f(\boldsymbol{A})=d(\boldsymbol{A})d(\boldsymbol{B})$, 即得公式 (5.15). □

5.12 非奇异矩阵的逆矩阵的行列式

由定义, 如果方阵 $\boldsymbol{A}$ 有一个左逆 $\boldsymbol{B}$ 使得 $\boldsymbol{BA}=\boldsymbol{I}$, 则称 $\boldsymbol{A}$ 为**非奇异的**. 若 $\boldsymbol{A}$ 存在左逆, 则左逆唯一且它同时也是右逆, 即 $\boldsymbol{AB}=\boldsymbol{I}$, 因此我们把 $\boldsymbol{B}$ 叫作 $\boldsymbol{A}$ 的**逆**, 记作 $\boldsymbol{A}^{-1}$, 因而 $\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}$. $\det\boldsymbol{A}$ 与 $\det(\boldsymbol{A}^{-1})$ 之间的关系如我们所想的那样自然.

定理 5.12 *若矩阵 $\boldsymbol{A}$ 非奇异, 则 $\det\boldsymbol{A}\neq 0$ 且*

$$\det(\boldsymbol{A}^{-1})=\frac{1}{\det\boldsymbol{A}}. \tag{5.16}$$

证明 由行列式乘积公式得 $(\det\boldsymbol{A})(\det(\boldsymbol{A}^{-1}))=\det(\boldsymbol{A}\boldsymbol{A}^{-1})=\det\boldsymbol{I}=1$. 因此 $\det\boldsymbol{A}\neq 0$ 并且等式 (5.16) 成立. □

由定理 (5.12) 知, $\det\boldsymbol{A}$ 不等于零是 $\boldsymbol{A}$ 非奇异矩阵的一个必要条件. 稍后 (定理 5.16) 我们将证明这个条件也是充分的, 即, 若 $\det\boldsymbol{A}\neq 0$, 则逆矩阵 $\boldsymbol{A}^{-1}$ 存在.

5.13 行列式与向量组的线性无关性

由定理 5.12 可以得到判别向量组线性无关性的一个准则.

定理 5.13 *n 个 n 维向量 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$ 线性无关的充分必要条件是行列式 $d(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n)$ 不等于零.*

证明 在定理 5.8 中我们已经证明, 若 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 线性相关, 则 $d(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n) = 0$. 为证明其逆, 假定 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 线性无关, 我们证明 $d(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n) \neq 0$.

令 V 为纯量的 n 序组组成的线性空间, 由于 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 是 n 维线性空间 V 中 n 个线性无关的向量, 因此它们组成 V 的一组基. 由定理 4.12, 存在将这 n 个向量映为单位坐标向量的一个线性变换 $T: V \to V$,

$$T(\boldsymbol{A}_k) = \boldsymbol{I}_k, \quad k = 1, 2, \cdots, n.$$

因此存在 $n \times n$ 矩阵 $\boldsymbol{B}$ 使得

$$\boldsymbol{A}_k \boldsymbol{B} = \boldsymbol{I}_k, \quad k = 1, 2, \cdots, n.$$

从而由引理 5.10, 我们有 $\boldsymbol{A}_k \boldsymbol{B} = (\boldsymbol{AB})_k$, 此处 $\boldsymbol{A}$ 是以 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n$ 为行的矩阵, 因此 $\boldsymbol{AB} = \boldsymbol{I}$, 于是 $\boldsymbol{A}$ 非奇异, 即 $\det \boldsymbol{A} \neq 0$. □

5.14 分块对角矩阵的行列式

设 $\boldsymbol{C}$ 为形如

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{bmatrix}$$

的方阵, 其中 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为方阵, $\boldsymbol{O}$ 表示零矩阵, 则 $\boldsymbol{C}$ 叫作以 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为主对角块的**分块** (block-diagonal) **对角矩阵**. 下述 5×5 矩阵

$$\boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 7 & 8 & 9 \end{bmatrix}$$

就是分块对角矩阵的一个例子, 它的主对角线上的两个块是

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

下述定理指出, 分块对角矩阵的行列式等于其主对角线上各块的行列式的乘积.

定理 5.14 *对任意两个方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$, 我们都有*

$$\det \begin{bmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{bmatrix} = (\det \boldsymbol{A})(\det \boldsymbol{B}). \tag{5.17}$$

证明 设 $\boldsymbol{A}$ 为一个 $n \times n$ 矩阵, $\boldsymbol{B}$ 为一个 $m \times m$ 矩阵, 则给定的分块对角矩阵可表示为下述两个分块对角矩阵的乘积:

$$\begin{bmatrix} A & O \\ O & B \end{bmatrix} = \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} \begin{bmatrix} I_n & O \\ O & B \end{bmatrix},$$

其中 I_n 与 I_m 分别表示 n 阶与 m 阶单位矩阵. 从而由行列式乘积公式得

$$\det \begin{bmatrix} A & O \\ O & B \end{bmatrix} = \det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} \det \begin{bmatrix} I_n & O \\ O & B \end{bmatrix}. \tag{5.18}$$

由于 $\begin{bmatrix} A & O \\ O & I_m \end{bmatrix}$ 的右上角的分块是一个零矩阵, 因此我们可以将行列式 $\det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix}$ 看作 A 的 n 个行向量的一个函数. 不难验证这个函数满足关于 n 阶行列式函数的所有 3 条公理, 从而由唯一性定理, 必有

$$\det \begin{bmatrix} A & O \\ O & I_m \end{bmatrix} = \det A.$$

同理可得 $\det \begin{bmatrix} I_n & O \\ O & B \end{bmatrix} = \det B$, 从而由等式 (5.18) 即得公式 (5.17). □

5.15 习　　题

1. 对下列关于 $n \times n$ 矩阵的各个命题, 或者证明它成立, 或者举出一个反例.
   (a) $\det(A + B) = \det A + \det B$.
   (b) $\det\{(A + B)^2\} = \{\det(A + B)\}^2$.
   (c) $\det\{(A + B)^2\} = \det(A^2 + 2AB + B^2)$.
   (d) $\det\{(A + B)^2\} = \det(A^2 + B^2)$.
2. (a) 将定理 5.14 推广到具有三个主对角块的分块对角矩阵的情形:

$$\det \begin{bmatrix} A & O & O \\ O & B & O \\ O & O & C \end{bmatrix} = (\det A)(\det B)(\det C).$$

   (b) 将定理 5.14 推广到具任意多个主对角块的情形并给出证明.
   (c) 证明: 若 A, B, C 为非奇异矩阵, 则分块对角矩阵 $\begin{bmatrix} A & O & O \\ O & B & O \\ O & O & C \end{bmatrix}$ 也是非奇异的并且它的逆矩阵是 $\begin{bmatrix} A^{-1} & O & O \\ O & B^{-1} & O \\ O & O & C^{-1} \end{bmatrix}$.
3. 设 $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & b & c & d \\ e & f & g & h \end{bmatrix}$, $B = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. 证明: $\det A = \begin{vmatrix} c & d \\ g & h \end{vmatrix}$, $\det B = \begin{vmatrix} a & b \\ e & f \end{vmatrix}$.

4. 将习题 3 的结论推广到 $n \times n$ 矩阵的情形并给出证明.

5. 设 $\boldsymbol{A} = \begin{bmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ e & f & g & h \\ x & y & z & w \end{bmatrix}$, 证明: $\det \boldsymbol{A} = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \det \begin{bmatrix} g & h \\ z & w \end{bmatrix}$.

6. 将习题 5 的结论推广到形如

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}$$

的 $n \times n$ 矩阵的情形并给出证明, 此处 $\boldsymbol{B}$, $\boldsymbol{C}$ 与 D 为方阵, $\boldsymbol{O}$ 表示零矩阵.

7. 用定理 5.13 确定下列向量组的线性相关性或线性无关性.

(a) $\boldsymbol{A}_1 = (1, -1, 0)$, $\boldsymbol{A}_2 = (0, 1, -1)$, $\boldsymbol{A}_3 = (2, 3, -1)$.

(b) $\boldsymbol{A}_1 = (1, -1, 2, 1)$, $\boldsymbol{A}_2 = (-1, 2, -1, 0)$, $\boldsymbol{A}_3 = (3, -1, 1, 0)$, $\boldsymbol{A}_4 = (1, 0, 0, 1)$.

(c) $\boldsymbol{A}_1 = (1,0,0,0,1)$, $\boldsymbol{A}_2 = (1,1,0,0,0)$, $\boldsymbol{A}_3 = (1,0,1,0,1)$, $\boldsymbol{A}_4 = (1,1,0,1,1)$, $\boldsymbol{A}_5 = (0,1,0,1,0)$.

5.16 行列式关于余子式的展开式

对于二阶和三阶行列式的情形, 在 5.10 节中, 我们通过将矩阵的各行表示单位坐标向量的线性组合来计算这个矩阵的行列式. 一般地, $n \times n$ 矩阵 $\boldsymbol{A}$ 的各行也都可以表示成 n 个单位坐标向量 $\boldsymbol{I}_1, \boldsymbol{I}_2, \cdots, \boldsymbol{I}_n$ 的线性组合, 例如, $\boldsymbol{A}$ 的第一行 $\boldsymbol{A}_1 = (a_{11}, a_{12}, \cdots, a_{1n})$ 可表示为

$$\boldsymbol{A}_1 = \sum_{j=1}^{n} a_{1j} \boldsymbol{I}_j$$

由于行列式关于第 1 行是线性的, 因此我们有

$$d(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n) = d\left(\sum_{j=1}^{n} a_{1j}\boldsymbol{I}_j, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n\right) = \sum_{j=1}^{n} a_{1j} d(\boldsymbol{I}_j, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n). \quad (5.19)$$

从而, 为计算 $\det \boldsymbol{A}$, 只需对各单位坐标向量 $\boldsymbol{I}_j$ 计算 $d(\boldsymbol{I}_j, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n)$ 即可. 我们用记号 $\boldsymbol{A}'_{1j}$ 表示将 $\boldsymbol{A}$ 中第 1 行 $\boldsymbol{A}_1$ 用单位坐标向量 $\boldsymbol{I}_j$ 代替所得到的矩阵. 例如, 当 $n = 3$ 时, 共有 3 个这样的矩阵:

$$\boldsymbol{A}'_{11} = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \boldsymbol{A}'_{12} = \begin{bmatrix} 0 & 1 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

$$\boldsymbol{A}'_{13} = \begin{bmatrix} 0 & 0 & 1 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

由 $\boldsymbol{A}'_{1j}$ 的意义可知 $\det \boldsymbol{A}'_{1j} = d(\boldsymbol{I}_j, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n)$, 因此展开式 (5.19) 可表示为下述形式:

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{1j} \det \boldsymbol{A}'_{1j}. \tag{5.20}$$

上式称为行列式的**展开式**(expansion formula), 它把 $\boldsymbol{A}$ 的行列式表示为 $\boldsymbol{A}$ 的第 1 行中各元素的线性组合, 数 $\det \boldsymbol{A}'_{1j}$ 叫作 a_{1j} 的**余子式**(cofactor), 记作 $\operatorname{cof} a_{1j}$. 因此展开式 (5.20) 又可表示为

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{1j} \operatorname{cof} a_{1j}, \tag{5.21}$$

叫作关于第 1 行余子式的展开式.

一般地, 我们可对任一元素 a_{kj} 定义它的余子式.

定义 (余子式)　*令 $\boldsymbol{A}'_{kj}$ 为将 $\boldsymbol{A}$ 中第 k 行 $\boldsymbol{A}_k$ 用单位坐标向量 $\boldsymbol{I}_j$ 代替所得到的矩阵, 则纯量 $\det \boldsymbol{A}'_{kj}$ 叫作元 a_{kj} 的余子式并记作 $\operatorname{cof} a_{kj}$, 于是我们有*

$$\operatorname{cof} a_{kj} = \det \boldsymbol{A}'_{kj}. \tag{5.22}$$

用于第 1 行的关于导出展开式 (5.21) 的论述也可用于第 k 行, 从而得到关于第 k 行的元素的展开式:

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{kj} \det \boldsymbol{A}'_{kj}. \tag{5.23}$$

由于 $\det \boldsymbol{A}'_{kj} = \operatorname{cof} a_{kj}$, 因此我们可将此结果表述为如下定理.

定理 5.15(行列式按第 k 行余子式的展开式)　*对任意 $n \times n$ 矩阵 $\boldsymbol{A}$ 和任意 $k = 1, 2, \cdots, n$, 我们都有*

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{kj}. \tag{5.24}$$

□

5.17　余子式矩阵

定理 5.12 告诉我们, 若方阵 $\boldsymbol{A}$ 非奇异, 则 $\det \boldsymbol{A} \neq 0$. 下面的定理 5.16 将证明它的逆定理也成立, 即若 $\det \boldsymbol{A} \neq 0$, 则 $\boldsymbol{A}^{-1}$ 存在. 不仅如此, 它还给出了用由 $\boldsymbol{A}$ 中各元素的代数余子式组成的矩阵来表示 $\boldsymbol{A}^{-1}$ 的公式, 这个用来表示 $\boldsymbol{A}^{-1}$ 的矩阵叫作 $\boldsymbol{A}$ 的余子式矩阵, 下面给出它的定义.

定义 (余子式矩阵)　给定 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$, 令 $\operatorname{cof} \boldsymbol{A}$ 表示以 $\operatorname{cof} a_{ij}$ 为 (i, j)-元的 $n \times n$ 矩阵, 则 $\operatorname{cof} \boldsymbol{A}$ 叫作 $\boldsymbol{A}$ 的余子式矩阵[①], 即

$$\operatorname{cof} \boldsymbol{A} = (\operatorname{cof} a_{ij})_{i,j=1}^{n}.$$

下述定理指出, $\boldsymbol{A}$ 与其余子式矩阵的转置矩阵的乘积是一个纯量与单位矩阵的乘积. (任意矩阵 $\boldsymbol{A}$ 的转置矩阵就是将 $\boldsymbol{A}$ 的行与列互换所得到的矩阵, 记作 $\boldsymbol{A}^{\mathrm{T}}$.

① 这里定义的余子式矩阵在通常的中文教材中称作伴随矩阵 (adjoint matrix). 在本书第 7 章中, "伴随矩阵" 一词用来表示一个矩阵的共轭转置矩阵.

转置矩阵的性质见 4.21 节的习题 7)

定理 5.16 设 $n \geqslant 2$, 对任意 $n \times n$ 矩阵 $\boldsymbol{A}$, 都有

$$\boldsymbol{A}(\operatorname{cof} \boldsymbol{A})^{\mathrm{T}} = (\det \boldsymbol{A})\boldsymbol{I}. \tag{5.25}$$

特别地, 如果 $\det \boldsymbol{A} \neq 0$, 则 $\boldsymbol{A}$ 的逆存在且由下式给出:

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}}(\operatorname{cof} \boldsymbol{A})^{\mathrm{T}}. \tag{5.26}$$

证明 我们利用定理 5.15 将 $\det \boldsymbol{A}$ 按它第 k 行的余子式展开:

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{kj}. \tag{5.27}$$

固定 k 并将上面的展开式用于一个新矩阵 $\boldsymbol{B}$, 这里 $\boldsymbol{B}$ 是对某个 $i \neq k$ 将 $\boldsymbol{A}$ 的第 i 行换为第 k 行而保持其余各行不变所得的矩阵. 由于 $\boldsymbol{B}$ 的第 i 行与 $\boldsymbol{A}$ 第 k 行相等, 因此 $\det \boldsymbol{B} = 0$, 将 $\det \boldsymbol{B}$ 按第 i 行的余子式展开, 则得

$$\det \boldsymbol{B} = \sum_{j=1}^{n} b_{ij} \operatorname{cof} b_{ij} = 0. \tag{5.28}$$

而由于 $\boldsymbol{B}$ 的第 i 行与 $\boldsymbol{A}$ 的第 k 行相等, 因此, 对每一个 j 都有

$$b_{ij} = a_{kj}, \quad \operatorname{cof} b_{ij} = \operatorname{cof} a_{ij}$$

于是由等式 (5.28) 得

$$\sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{ij} = 0, \quad k \neq i. \tag{5.29}$$

由于 $k \neq i$, 因此展开式 (5.29) 常叫作**按相异余子式展开** (expansion by alien cofactors). (5.28) 与 (5.29) 可结合起来表述为下式:

$$\sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{ij} = \begin{cases} \det \boldsymbol{A}, & i = k \\ 0, & i \neq k. \end{cases} \tag{5.30}$$

由于等式 (5.30) 左边是矩阵乘积 $\boldsymbol{A}(\operatorname{cof} \boldsymbol{A})^{\mathrm{T}}$ 的 (k, i)-元, 因此由等式 (5.30) 即得公式 (5.25). □

作为定理 5.12 与定理 5.16 的直接推论, 我们得到关于方阵为非奇异的下述充分必要条件.

定理 5.17 方阵 $\boldsymbol{A}$ 可逆的充分必要条件是 $\det \boldsymbol{A} \neq 0$. □

5.18 Cramer 法则

定理 5.16 可用来给出以非奇异矩阵为系数矩阵的线性方程组解的明确表达式. 为纪念瑞士数学家 Gabriel Cramer(1704—1752), 将这个公式叫作 **Cramer 法则**. $n = 3$ 的情形已在 2.14 节中讨论过.

定理 5.18 (Cramer 法则) 设 n 个关于 n 个未知量 $x_1, \cdots, x_n$ 的方程组成的线性方程组

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \quad (i = 1, 2, \cdots, n)$$

的系数矩阵 $\boldsymbol{A} = (a_{ij})$ 为非奇异矩阵，则这个方程组有唯一的一组解，并且这组解由下面公式给出：

$$x_j = \frac{1}{\det \boldsymbol{A}} \sum_{k=1}^{n} b_k \operatorname{cof} a_{kj}, \quad j = 1, 2, \cdots, n. \tag{5.31}$$

证明 这个方程组可写成矩阵方程的形式，$\boldsymbol{A}X = \boldsymbol{B}$，其中 X 与 $\boldsymbol{B}$ 都是 n 维列向量，$\boldsymbol{X} = [x_1, \cdots, x_n]^{\mathrm{T}}$，$\boldsymbol{B} = [b_1, \cdots, b_n]^{\mathrm{T}}$. 由于 $\boldsymbol{A}$ 非奇异，因此这个矩阵方程有唯一解，且此解是

$$X = \boldsymbol{A}^{-1}\boldsymbol{B} = \frac{1}{\det \boldsymbol{A}} (\operatorname{cof} \boldsymbol{A})^{\mathrm{T}} \boldsymbol{B}. \tag{5.32}$$

比较等式 (5.32) 中的各个分量即得公式 (5.31). □

我们指出，若令 $\boldsymbol{C}_j$ 表示将系数矩阵 $\boldsymbol{A}$ 中的第 j 列换为列矩阵 $\boldsymbol{B}$ 而保持其余各列不变所得的矩阵，则公式 (5.31) 中的求和式就是矩阵 $\boldsymbol{C}_j$ 的行列式. 因此，公式 (5.31) 中的 x_j 也可表示为两个行列式的商的形式：

$$x_j = \frac{\det \boldsymbol{C}_j}{\det \boldsymbol{A}}.$$

5.19 行列式按子式的展开式

n 阶方阵 $\boldsymbol{A}$ 的行列式可按其第 k 行元素的余子式展开如下：

$$\det \boldsymbol{A} = \sum_{j=1}^{n} a_{kj} \operatorname{cof} a_{kj},$$

此处 $\operatorname{cof} a_{kj}$ 是将 $\boldsymbol{A}$ 中的第 k 个行用单位坐标向量 $\boldsymbol{I}_j$ 代替而保持 $\boldsymbol{A}$ 中其余各行不变所得的矩阵 $\boldsymbol{A}'_{kj}$ 的行列式. 本节将证明，除去可能相差一个正负号之外，$\det \boldsymbol{A}'_{kj}$ 等于一个 $n-1$ 阶方阵的行列式. 这些 $n-1$ 阶行列式叫作**子式** (minor).

定义 (子式) 设 $n \geqslant 2$，$\boldsymbol{A}$ 为一个给定的 n 阶方阵，从 $\boldsymbol{A}$ 中去掉第 k 行和第 j 列后剩下的 $n-1$ 阶方阵叫作 $\boldsymbol{A}$ 的 (k, j)-子式，记作 $\boldsymbol{A}_{kj}$.

例 一个三阶方阵 $\boldsymbol{A} = (a_{ij})$ 共有 9 个子式，每个元素都对应一个子式. 第一行各元素的子式为

$$\boldsymbol{A}_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad \boldsymbol{A}_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad \boldsymbol{A}_{13} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

公式 (5.2) 将一个 3×3 矩阵的行列式表示为第一行各子式的行列式的线性组合，公式为

$$\det \boldsymbol{A} = a_{11} \det \boldsymbol{A}_{11} - a_{12} \det \boldsymbol{A}_{12} + a_{13} \det \boldsymbol{A}_{13}.$$

下面的定理将这个展开式推广到任意 $n \times n$ 矩阵的情形. □

定理 5.19 (行列式按第 k 行子式的展开) 设 $n \geqslant 2$, $\boldsymbol{A}$ 为 $n \times n$ 矩阵, 则 a_{kj} 的余子式 $\operatorname{cof} a_{kj}$ 与子式 $\boldsymbol{A}_{kj}$ 之间有如下关系:

$$\det \boldsymbol{A}'_{kj} = (-1)^{k+j} \det \boldsymbol{A}_{kj}, \tag{5.33}$$

从而 $\det \boldsymbol{A}$ 关于第 k 行子式的展开式由下述公式给出:

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \boldsymbol{A}_{kj}. \tag{5.34}$$

证明 为阐述证明思想, 我们先考虑 $k = j = 1$ 时的特殊情形. 矩阵 $\boldsymbol{A}'_{11}$ 为下述矩阵:

$$\boldsymbol{A}'_{11} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}.$$

通过第三类行初等变换可将第一列中位于 1 之下的各分量都化为零而保持其余分量不变. 例如, 如果我们将 $\boldsymbol{A}'_{11}$ 的第一行乘以 $-a_{21}$ 之后加到第二行, 则新的第二行为 $(0, a_{22}, a_{23}, \cdots, a_{2n})$. 反复运用这样的行初等变换得到的新矩阵的第一列为 $(1, 0, 0, \cdots, 0)^{\mathrm{T}}$ 而其余各列则与 $\boldsymbol{A}'_{11}$ 相同. 这个新矩阵是一个分块对角阵, 因此由定理 5.14, 它的行列式等于 $\boldsymbol{A}_{11}$ 此处 $\boldsymbol{A}_{11}$ 是 $\boldsymbol{A}$ 的 $(1,1)$-子式:

$$\boldsymbol{A}_{11} = \begin{bmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}.$$

因此 $\det \boldsymbol{A}'_{11} = \det \boldsymbol{A}_{11}$, 由此证明当 $k = j = 1$ 时等式 (5.33) 成立.

接下来我们考虑 $k = 1$ 而 j 任意的情形并证明

$$\det \boldsymbol{A}'_{1j} = (-1)^{j-1} \det \boldsymbol{A}_{1j}. \tag{5.35}$$

一旦我们证明等式 (5.35) 成立, 即可立刻推出更一般的等式 (5.33), 这是因为矩阵 $\boldsymbol{A}'_{kj}$ 可以经 $k-1$ 次交换相邻两行化为一个矩阵 $\boldsymbol{B}'_{1j}$. 由于每经过这样一次交换, 行列式都改变一次正负号, 因此得,

$$\det \boldsymbol{A}'_{kj} = (-1)^{k-1} \det \boldsymbol{B}'_{1j}, \tag{5.36}$$

此处 $\boldsymbol{B}$ 是第一行为 $\boldsymbol{I}_j$, $(1,j)$-子式 $\boldsymbol{B}_{1j} = \boldsymbol{A}_{kj}$ 的 $n \times n$ 矩阵, 由等式 (5.35) 得

$$\det \boldsymbol{B}'_{1j} = (-1)^{j-1} \det \boldsymbol{B}_{1j} = (-1)^{j-1} \det \boldsymbol{A}_{kj},$$

因此由等式 (5.36) 得

$$\det \boldsymbol{A}'_{kj} = (-1)^{k-1}(-1)^{j-1} \det \boldsymbol{A}_{kj} = (-1)^{k+j} \det \boldsymbol{A}_{kj}.$$

因此若我们证得等式 (5.35), 也就证明了等式 (5.33).

现在我们来证明等式 (5.35), 矩阵 $\boldsymbol{A}'_{1j}$ 有如下形式:

$$\boldsymbol{A}'_{1j} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3j} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}.$$

通过第三类行初等变换可将 $\boldsymbol{A}'_{1j}$ 的第 j 列中 1 以下的元素都化为零, 用 $\boldsymbol{A}^0_{1j}$ 表示所得的矩阵, 即

$$\boldsymbol{A}^0_{1j} = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ a_{31} & \cdots & 0 & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{bmatrix}.$$

由于矩阵经过第三类行初等变换其行列式不变, 因此 $\det \boldsymbol{A}^0_{1j} = \det \boldsymbol{A}'_{1j}$. $\boldsymbol{A}$ 的 $(1,j)$-子式 $\boldsymbol{A}_{1j}$ 的形式为

$$\boldsymbol{A}_{1j} = \begin{bmatrix} a_{21} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3j-1} & a_{3j+1} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{bmatrix}.$$

现在我们把 $\boldsymbol{A}^0_{1j}$ 看作 $\boldsymbol{A}_{1j}$ 的 $n-1$ 个行的一个函数, 设为 $\det \boldsymbol{A}^0_{1j} = f(\boldsymbol{A}_{1j})$, 则 f 满足关于 $n-1$ 阶行列式函数的前两条公理. 从而由唯一性定理得

$$f(\boldsymbol{A}_{1j}) = f(J) \det \boldsymbol{A}_{1j},$$

其中 J 是 $n-1$ 阶单位矩阵. 于是, 为证明等式 (5.35), 我们必须证明 $f(J) = (-1)^{j-1}$. 令 $\boldsymbol{C}$ 表示下述 $n \times n$ 矩阵:

$$\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ 0 & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \leftarrow \text{第 } j \text{ 行}$$

第 j 列

位于斜的虚线上的所有元素都是 1, 其余未标明的元素都是零. 由 f 的定义得 $f(J) = \det \boldsymbol{C}$. 将 $\boldsymbol{C}$ 中的第 1 行依次与第 $2,3,\cdots,j$ 行相交换, 经过这 $j-1$ 次交换得到 $n \times n$ 单位矩阵. 由于每经过一次上述行交换行列式都改变一次正负号, 因此 $\det \boldsymbol{C} = (-1)^{j-1}$, 从而又得 $f(J) = (-1)^{j-1}$, 由此即得等式 (5.35), 于是得等式 (5.33). □

5.20 习　　题

1. 求下列各矩阵的余子式矩阵.

(a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.　(b) $\begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 1 \\ -1 & -2 & 0 \end{bmatrix}$.　(c) $\begin{bmatrix} 3 & 1 & 2 & 4 \\ 2 & 0 & 5 & 1 \\ 1 & -1 & -2 & 6 \\ -2 & 3 & 2 & 3 \end{bmatrix}$.

2. 求习题 1 中各非奇异矩阵的逆矩阵.
3. 求出所有使 $\lambda \boldsymbol{I} - \boldsymbol{A}$ 为奇异矩阵的纯量 λ, $\boldsymbol{A}$ 为:

(a) $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,　(b) $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix}$,　(c) $\boldsymbol{A} = \begin{bmatrix} 11 & -2 & 8 \\ 19 & -3 & 14 \\ -8 & 2 & -5 \end{bmatrix}$.

4. 设 $n \geqslant 2$, $\boldsymbol{A}$ 为 $n \times n$ 矩阵, 证明关于余子式矩阵的下列性质.
   (a) $\operatorname{cof}(\boldsymbol{A}^{\mathrm{T}}) = (\operatorname{cof} \boldsymbol{A})^{\mathrm{T}}$.
   (b) $(\operatorname{cof} \boldsymbol{A})^{\mathrm{T}} \boldsymbol{A} = (\det \boldsymbol{A}) \boldsymbol{I}$.
   (c) $\boldsymbol{A}(\operatorname{cof} \boldsymbol{A})^{\mathrm{T}} = (\operatorname{cof} \boldsymbol{A})^{\mathrm{T}} \boldsymbol{A}$. (即 $\boldsymbol{A}$ 与其余子式矩阵的转置阵可交换.)
5. 用 Cramer 法则解下列线性方程组.
   (a) $x + 2y + 3z = 8$, $2x - y + 4z = 7$, $-y + z = 1$.
   (b) $x + 2y + 2z = 0$, $3x - y - z = 3$, $2x + 5y + 3z = 4$.

5.21 行列式函数的存在性

在本节中, 我们通过对方阵的阶数作数学归纳法来证明: 对每一个 n, n 阶的行列式函数都存在. 对 $n = 1,2,3$ 的情形, 我们已经证明了其存在性.

假定 $n-1$ 阶行列式函数存在, 从逻辑上来说, 将某个关于子式的展开式, 例如关于第一行子式的展开式, 用作 n 阶行列式函数的备选对象似乎是合理的. 不过, 如果我们用关于第一列子式的类似公式将更容易验证各公理.

定理 5.20　*假定 $n-1$ 行列式函数存在, 对 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$, 令 f 为由下式定义的一个函数:*

$$f(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_n) = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} \det \boldsymbol{A}_{j1}. \tag{5.37}$$

则 f 满足关于 n 阶行列式函数的所有三条公理, 从而对每一个 n, n 阶行列式函数都存在.

证明 由于 $\boldsymbol{A}_{j1}$ 是 $\boldsymbol{A}$ 的 $(j,1)$-子式, 它包含了 $\boldsymbol{A}$ 中 $n-1$ 个去掉第一个元素的行 ($\boldsymbol{A}$ 中除 $\boldsymbol{A}_j$ 之外的所有行, 并将每一行中的第一个元素去掉之后作为 $\boldsymbol{A}_{j1}$ 的行). 由归纳假定, $\det \boldsymbol{A}_{j1}$ 满足关于 $n-1$ 阶行列式函数的所有公理. 下面我们验证由展开式 (5.37) 定义的函数 f 也满足关于 n 阶行列式函数的全部三条公理.

先考虑公理 1, 即齐性. 将展开式 (5.37) 表成下列形式:

$$f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)=\sum_{j=1}^{n} f_j(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n),$$

此处

$$f_j(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)=(-1)^{j+1}a_{j1}\det \boldsymbol{A}_{j1}.$$

我们要证明每个 f_j 都满足齐性, 从而 f 亦然.

我们先来研究将 $\boldsymbol{A}$ 的第一行乘以纯量 t 以后所产生的影响. 由于 $\boldsymbol{A}_{11}$ 不包含第一行 $\boldsymbol{A}_1$, 因此 $\boldsymbol{A}_{11}$ 不受影响, 但是其系数 a_{11} 被 t 乘了, 所以我们有

$$f_j(t\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)=tf_j(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n).$$

因此各个 f_j 关于第一行都是齐次的.

今设 $k>1$, 若 $\boldsymbol{A}$ 的第 k 行乘以 t, 则子式 $\boldsymbol{A}_{k1}$ 不受 t 影响, 但其系数 a_{k1} 也乘以 t, 所以 f_k 关于第 k 行也是齐次的. 若 $j\neq k$, 则系数 a_{j1} 不受影响, 但 $\boldsymbol{A}_{j1}$ 中恰有一行乘以 t, 因此关于 f_j 第 k 行也是齐次的. 从而每一个 f_j 关于第 k 行都是齐次的, 所以 f 满足公理 1.

现在我们证明 f 满足公理 2, 即行相加不变性. 在这个情形, 在和式的各项 f_j 中正好只有两项受到影响而其余各项都未受影响, 并且受到影响的两项的和并未改变, 因此整个的和式也未改变. 为说明这点, 令 $\boldsymbol{B}$ 表示从 $\boldsymbol{A}$ 将行 $\boldsymbol{A}_1$ 加到 $\boldsymbol{A}_k$ 之后所得的矩阵, 此处 $k\neq 1$, 即对所有 $i\neq k$,

$$\boldsymbol{B}_k=\boldsymbol{A}_1+\boldsymbol{A}_k,\quad \boldsymbol{B}_i=\boldsymbol{A}_i.$$

从而对所有 $i\neq k$ 及所有 j,

$$b_{kj}=a_{1j}+a_{kj},\quad b_{ij}=a_{ij}.$$

在定义 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$ 的和式 (5.37) 中, 只有 $j=1$ 与 $j=k$ 的两项才受到影响. 例如, 若取 $k=2$, 则

$$\boldsymbol{B}_2=\boldsymbol{A}_1+\boldsymbol{A}_2,\quad b_{2j}=a_{1j}+a_{2j}.$$

在定义 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$ 的和式 (5.37) 中 $j=1$ 的项是

$$\begin{aligned}b_{11}\det \boldsymbol{B}_{11}&=b_{11}d(\boldsymbol{B}_2,\boldsymbol{B}_3,\cdots,\boldsymbol{B}_n)=a_{11}d(\boldsymbol{A}_1+\boldsymbol{A}_2,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)\\&=a_{11}d(\boldsymbol{A}_1,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)+a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n).\end{aligned}$$

$j=2$ 的项为

$$\begin{aligned}-b_{21}\det \boldsymbol{B}_{21}&=-b_{21}d(\boldsymbol{B}_1,\boldsymbol{B}_3,\cdots,\boldsymbol{B}_n)=-(a_{11}+a_{21})d(\boldsymbol{A}_1,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)\\&=-a_{11}d(\boldsymbol{A}_1,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)-a_{21}d(\boldsymbol{A}_1,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n).\end{aligned}$$

将 $j=1$ 与 $j=2$ 的两项相加并由消去律得

$$\begin{aligned} b_{11}\det \boldsymbol{B}_{11}-b_{21}\det \boldsymbol{B}_{21} &= a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)-a_{21}d(\boldsymbol{A}_1,\boldsymbol{A}_3,\cdots,\boldsymbol{A}_n)\\ &= a_{11}\det \boldsymbol{A}_{11}-a_{21}\det \boldsymbol{A}_{21}. \end{aligned}$$

由于其余各项都未变, 因此得 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$.

类似地, 若取 $k=3$, 则有

$$\boldsymbol{B}_3=\boldsymbol{A}_1+\boldsymbol{A}_3,\quad b_{3j}=a_{1j}+a_{3j}.$$

在定义 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)$ 的和式 (5.37) 中 $j=1$ 的项为

$$\begin{aligned} b_{11}\det \boldsymbol{B}_{11}=b_{11}d(\boldsymbol{B}_2,\boldsymbol{B}_3,\cdots,\boldsymbol{B}_n) &= a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_1+\boldsymbol{A}_3,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)\\ &= a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_1,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)+a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_3,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n), \end{aligned}$$

$j=3$ 的项为

$$\begin{aligned} b_{31}\det \boldsymbol{B}_{31}=b_{31}d(\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_4,\cdots,\boldsymbol{B}_n) &= (a_{11}+a_{31})d(\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)\\ &= a_{11}d(\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)+a_{31}d(\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)\\ &= -a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_1,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)+a_{31}d(\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n). \end{aligned}$$

将 $j=1$ 与 $j=3$ 的两项相加并由消去律得

$$\begin{aligned} b_{11}\det \boldsymbol{B}_{11}+b_{31}\det \boldsymbol{B}_{31} &= a_{11}d(\boldsymbol{A}_2,\boldsymbol{A}_3,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)+a_{31}d(\boldsymbol{A}_1,\boldsymbol{A}_2,\boldsymbol{A}_4,\cdots,\boldsymbol{A}_n)\\ &= a_{11}\det \boldsymbol{A}_{11}+a_{31}\det \boldsymbol{A}_{31}. \end{aligned}$$

由于其余项都不变, 因而在此情形也有 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$. 同理可证, 在一般情形时都有 $f(\boldsymbol{B}_1,\cdots,\boldsymbol{B}_n)=f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)$, 因此 f 满足公理 2.

最后, 我们来验证 f 满足公理 3. 当 $\boldsymbol{A}=\boldsymbol{I}$ 时, 我们有 $a_{11}=1$ 而对所有的 $j>1$ 都有 $a_{j1}=0$, 而且子式 $\boldsymbol{A}_{11}$ 是 $n-1$ 阶单位矩阵, 因此在和式 (5.37) 中, 除了第 1 项等于 1 之外, 其余各项都是零, 从而 $f(\boldsymbol{I}_1,\cdots,\boldsymbol{I}_n)=1$, 即 f 满足公理 3. □

在前面的论证中, 我们完全可以用第 k 列子式 $\boldsymbol{A}_{jk}$ 取代第 1 列子式 $\boldsymbol{A}_{j1}$ 来定义一个函数 f. 事实上, 若令

$$f(\boldsymbol{A}_1,\cdots,\boldsymbol{A}_n)=\sum_{j=1}^{n}(-1)^{j+k}a_{jk}\det \boldsymbol{A}_{jk}, \tag{5.38}$$

则由完全一样的证明可知此 f 也满足关于行列式函数的所有三条公理. 再由唯一性定理可知这个关于第 k 列子式的展开式等于 $\det \boldsymbol{A}$.

展开公式 (5.38) 不但建立了行列式函数的存在性, 而且也揭示了行列式的行的性质与列的性质之间的一种联系. 这一联系在下面的定理中作进一步讨论.

定理 5.21 *方阵和它的转置矩阵有相同行列式.*

证明 对方阵的阶数 n 作数学归纳法. 当 $n=1$ 时结论显然. 当 $n=2$ 时, 结论可简单地由等式 (5.1) 导出. 今设定理对 $n-1$ 阶方阵都成立, 令 $\boldsymbol{A}=(a_{ij})$, 再令 $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{T}}=(b_{ij})$ 为 $\boldsymbol{A}$ 的转置矩阵. 将 $\det \boldsymbol{A}$ 用它的第 1 列子式展开, 再将 $\det \boldsymbol{B}$ 用它的第 1 行子式展开, 则得

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} \det \boldsymbol{A}_{j1}, \quad \det \boldsymbol{B} = \sum_{j=1}^{n} (-1)^{j+1} b_{1j} \det \boldsymbol{B}_{1j},$$

但由转置矩阵的定义, 我们有 $b_{1j} = a_{j1}$ 及 $\boldsymbol{B}_{1j} = (\boldsymbol{A}_{j1})^{\mathrm{T}}$. 由于我们假设定理对 $n-1$ 阶方阵成立, 因此有 $\det \boldsymbol{B}_{1j} = \det \boldsymbol{A}_{j1}$, 于是上面关于 $\det \boldsymbol{A}$ 与 $\det \boldsymbol{B}$ 的和式中的各项都对应相等, 从而得 $\det \boldsymbol{A} = \det \boldsymbol{B}$. □

5.22 关于行列式的综合性习题

1. (a) 说明为什么下面两式都是 xy 平面上过不同两点 (x_1, y_1) 与 (x_2, y_2) 的直线的笛卡儿方程:

$$\det \begin{bmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{bmatrix} = 0; \quad \det \begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = 0.$$

   (b) 对 $\mathbb{R}^3$ 中过三个不同点的平面叙述并证明相应的结论.

   (c) 对 xy 平面上过不共线三点的圆叙述并证明相应结论.

2. 本题给出构造逆为整数矩阵的非奇异整数矩阵的一个有用的方法. 令 $\boldsymbol{A} = \begin{bmatrix} 1 & x & y \\ 0 & 1 & x \\ 0 & 0 & 1 \end{bmatrix}$.

   (a) 求 $\boldsymbol{A}$ 的余子式矩阵.

   (b) 令 $\boldsymbol{A}^{\mathrm{T}}$ 表示 $\boldsymbol{A}$ 的转置矩阵, $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$, 证明 $\boldsymbol{B}$ 为非奇异矩阵.

   (c) 若 x 与 y 都是整数, 证明 $\boldsymbol{B}$ 与 $\boldsymbol{B}^{-1}$ 都是整数矩阵.

3. 下述行列式叫作 $n \times n$ Vandermonde 矩阵的行列式或 n 阶 Vandermonde 行列式:

$$V_n(a_1, a_2, \cdots, a_n) = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}.$$

   其中 $a_1, a_2, \cdots, a_n$ 可以为实数或复数, 若诸 a_i 中有两个相等, 则由于有两行相同, 因此行列式等于零. 设各 a_i 互不相同, 并用 x 代替 a_1, 即令 $P(x) = V_n(x, a_2, \cdots, a_n)$, 则 $P(x)$ 的第一行为 $1, x, x^2, \cdots, x^{n-1}$.

   (a) 证明: $P(x)$ 是 x 的次数 $\leqslant n-1$ 的多项式. [提示: 将行列式按第 1 行子式展开.]

   (b) 证明: $P(x)$ 有 $n-1$ 个不同的根 $a_2, \cdots, a_n$, 因此有分解式:

$$P(x) = k \prod_{i=2}^{n} (x - a_i),$$

   此处常数因子 k 是 x^{n-1} 的系数.

   (c) 证明: $k = (-1)^{n-1} V_{n-1}(a_2, \cdots, a_n)$.

   (d) 利用 (b) 与 (c) 推导下述递推公式:

$$V_n(a_1, a_2, \cdots, a_n) = \prod_{i=2}^{n} (a_i - a_1) V_{n-1}(a_2, \cdots, a_n).$$

   (e) 利用 (d) 中的递推关系证明:

$$V_n(a_1, a_2, \cdots, a_n) = \prod_{\substack{i,j=1 \\ i<j}}^{n} (a_j - a_i).$$

上式指出, 若各个 a_i 互不相同, 则 Vandermonde 行列式不等于零, 即 Vandermonde 矩阵为非奇异阵, 这一性质有广泛应用.

习题 4~7 要用到微分学的知识.

4. 给定 n^2 个在区间 (a,b) 上都可微的函数 f_{ij}, 对 (a,b) 中的每个 x, 定义 $F(x)=\det[f_{ij}(x)]$. 证明: 导数 $F'(x)$ 是下列 n 个行列式之和:
$$F'(x)=\sum_{i=1}^{n}\det \boldsymbol{A}_i(x),$$
此处 $\boldsymbol{A}_i(x)$ 是对矩阵 $[f_{ij}(x)]$ 的第 i 行中各函数求导后所得的矩阵.

5. 设 $F(x)=\det\begin{bmatrix}1 & x & x^2 & x^3\\ 0 & 1 & 2x & 3x\\ 0 & 0 & 2 & 3\\ 1 & \mathrm{e}^x & \mathrm{e}^{2x} & \mathrm{e}^{3x}\end{bmatrix}$, 求导数 $F'(x)$.

6. 若 $F(x)=\det\begin{bmatrix}x & \dfrac{x^2}{2} & \dfrac{x^3}{3} & \dfrac{x^4}{4} & \dfrac{x^5}{5}\\ 1 & 1 & 1 & 1 & 1\\ 1 & 2 & 4 & 8 & 16\\ 1 & 3 & 9 & 27 & 81\\ 1 & 4 & 16 & 64 & 256\end{bmatrix}$, 证明: 存在常数 a,b,c, 使得 $F'(x)=c(x^2-3x+2)(x^2+ax+b)$.

7. 设 $W(x)=[u_j^{(i-1)}(x)]$ 为一个 $n\times n$ 函数矩阵, 其中第一行以下各行都为其上面一行的各元素求导后所得, 为纪念波兰数学家 J. M. H. Wronski, 称 $W(x)$ 为一个 **Wronski 矩阵** (我们在第 8 章中将用到这个矩阵). 证明: $\det W(x)$ 的导数是将 $W(x)$ 最后一行各元素求导后所得到的矩阵的行列式. [提示: 应用习题 4 的结果.]

第 6 章　特征值与特征向量

6.1　具有对角矩阵表示的线性变换

令 $T:V\to V$ 为作用在有限维线性空间 V 上的一个线性变换. T 的与 V 的坐标系 (基) 无关的性质叫作 T 的内蕴 (intrinsic) 性质. T 的内蕴性质为 T 的全体矩阵表示所共有. 如果能选取一组基使所得到的矩阵表示的形式特别简单, 则有可能由此矩阵表示直接推出线性变换 T 的某些内蕴性质.

对角矩阵是一类特别简单的矩阵, 因此我们要问, 是否每一个线性变换都有对角矩阵表示? 在第 4 章中, 我们讨论了寻求线性变换 $T:V\to W$ 的对角矩阵表示的问题, 此处 $\dim V=n$, $\dim W=m$. 定理 4.14 指出, 我们总能找到 V 的一组基 $(v_1,\cdots,v_n)$ 与 W 的一组基 $(w_1,\cdots,w_n)$ 使得 T 关于这样选取的两组基的矩阵是一个对角矩阵. 特别地, 若 $V=W$, 则所得到的矩阵是一个对角矩阵. 但如果我们要求 V 与 W 用的是同一组基, 则情形又将如何? 由此便引出本章的中心问题: 给定有限维线性空间 V 的一个线性变换 $T:V\to V$, 是否存在 V 的一组基使得 T 在这组基下的矩阵为对角矩阵?

我们采用下述记号: 若 $\boldsymbol{A}=(a_{ij})$ 是一个对角矩阵, 则记作 $\boldsymbol{A}=\operatorname{diag}(a_{11}, a_{22},\cdots,a_{nn})$.

本章的第一个定理揭示了这样一个事实: 一个线性变换由它在基元素上的作用所完全确定.

定理 6.1　设 $\dim V=n$, 给定线性变换 $T:V\to V$, 则 T 在 V 的某组基 $(v_1,\cdots,v_n)$ 下的矩阵为对角矩阵的充分必要条件是存在纯量 $\lambda_1,\cdots,\lambda_n$ 使得

$$T(v_k)=\lambda_k v_k,\quad k=1,2,\ldots,n, \tag{6.1}$$

此时, T 在基 $(v_1,\cdots,v_n)$ 下的矩阵是 $\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$.

证明　若 T 在基 $(v_1,\cdots,v_n)$ 下的矩阵是对角矩阵 $\boldsymbol{A}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, 则 T 在基元素上的作用由式 (6.1) 给出. 反之, 若存在纯量 $\lambda_1,\cdots,\lambda_n$, 使得对 $1\leqslant k\leqslant n$, 基元素 v_k 被映为 $\lambda_k v_k$, 则 $\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$ 便是 T 在基 $(v_1,\cdots,v_n)$ 下的矩阵. □

于是, 寻找线性变换的对角矩阵表示的问题便等价于寻求 V 中 n 个线性无关的元素 $v_1,\cdots,v_n$ 和 n 个纯量 $\lambda_1,\cdots,\lambda_n$ 使等式 (6.1) 对每一个基元素 v_k 都成立的问题.

满足等式 (6.1) 的纯量 λ_k 和非零元素 v_k 分别叫作线性变换 T 的特征值 (eigenvalue) 与特征向量 (eigenvector)①. 下一节中, 我们将在一般情形研究特征值与特征向量.

6.2 线性变换的特征值与特征向量

我们用 V 表示线性空间, S 表示 V 的子空间, 空间 V 与子空间 S 并不要求是有限维的.

定义 (特征值与特征向量) *令 $T: S \to V$ 为 S 到 V 中的一个线性变换. 设 λ 为一个纯量, 若存在 S 中的一个非零元素 x 使得*

$$T(x) = \lambda x \tag{6.2}$$

*则称 λ 为 T 的一个***特征值***, 并称 x 为 T 的属于 λ 的一个***特征向量***, λ 叫对应于 x 的特征值.*

对给定的特征向量 x, 只有唯一的一个特征值与之对应. 事实上, 若对某个 $x \neq O$ 同时有 $T(x) = \lambda x$ 与 $T(x) = \mu x$, 则 $\lambda x = \mu x$, 从而由 $x \neq O$, 得 $\lambda = \mu$.

若 $V = S = \mathbb{R}^n$, 则当 $\lambda \neq 0$ 时, 等式 $T(x) = \lambda x$ 有一个简单的几何解释: 每一个特征向量 x 都被映为一个与 x 平行的向量.

注 尽管等式 (6.2) 对 $x = O$ 及任意一个纯量 λ 都成立, 但在定义中并不将零元素看作特征向量. 这样做的理由之一是要保证与给定特征向量对应的特征值唯一.

下述各例用来说明上述概念的意义.

例 1 (纯量乘法) 令 $T: S \to V$ 为由等式 $T(x) = cx$ 所定义的线性变换, 此处 x 为 S 中的任意元素, c 是一个固定的纯量. 在本例中, S 中的任一非零元素都是属于特征值 c 的特征向量. □

例 2 (由所有使得 $T(x) = \lambda x$ 的 x 组成的特征空间 $E(\lambda)$) 设 $T: S \to V$ 为一个线性变换, λ 为 T 的一个特征值. 令 $E(\lambda)$ 表示 S 中所有满足条件 $T(x) = \lambda x$ 的元素 x 的集合, 则 $E(\lambda)$ 包含零元素 O 和所有属于 λ 的特征向量. 若 $x, y \in E(\lambda)$, 则对所有的纯量 a 与 b, 都有

$$T(ax + by) = aT(x) + bT(y) = a\lambda x + b\lambda y = \lambda(ax + by),$$

因此 $ax+by \in E(\lambda)$. 从而 $E(\lambda)$ 是 S 的一个子空间, 叫作 λ 的**特征空间** (eigenspace). $E(\lambda)$ 可能是有限维的, 也可能是无限维的. 当 $E(\lambda)$ 为有限维时, 由于它至少包含一个属于 λ 的非零元素 (特征向量) x, 因此 $E(\lambda)$ 的的维数至少是 1. □

---

① 单词 eigenvector 和 eigenvalue 分别是德语单词 Eigenvektor 和 Eigenwert 的英译. 用些作者用术语 characteristic vector 或 proper vector 作为 eigenvector 的同义语. 特征值也被称为 characteristic value, proper value 或 latent root.

例 3 (零特征值的存在性) 由定义, 特征向量若存在, 则它必不可能为零元素. 然而, 纯量 0 则有可能是特征值. 事实上, 若 0 是对应于特征向量 x 的特征值, 则 $T(x) = 0x = O$, 即 x 在 T 的**零空间**中. 反之, 若 T 的零空间包含有非零元素, 则这些非零元素都是属于特征值 0 的特征元素. 对于一个一般的特征值 λ, 不论 λ 是否为零, $E(\lambda)$ 都是线性变换 $T - \lambda I$ 的零空间. □

例 4 (xy 平面的反射变换) 设 $S = V = \mathbb{R}^3$, 令 T 为 xy 平面的**反射变换** (reflection), 即 T 是下列方式作用在基向量 $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ 上的线性变换:

$$T(\boldsymbol{i}) = \boldsymbol{i}, T(\boldsymbol{j}) = \boldsymbol{j}, T(\boldsymbol{k}) = -\boldsymbol{k}.$$

则 xy 平面上的任一非零向量都是属于特征值 1 的特征向量. 其余的任一特征向量都可以表示成 $c\boldsymbol{k}$ 的形式, 此处 $c \neq 0$. 它们都是属于特征值 -1 的特征向量. □

例 5 (平面转动一个固定角度 α 的旋转变换) 这个例子特别有意思, 因为它说明特征向量的存在性与纯量域有关. 我们可以用两种方法将 xy 平面看作线性空间: (1) 二维**实**线性空间, $V = \mathbb{R}^2$, 它的基向量为 $\boldsymbol{i} = (1,0)$ 与 $\boldsymbol{j} = (0,1)$, 以实数为纯量; (2) 一维**复**线性空间, $V = \mathbb{C}$, 基向量为 1, 以复数为纯量.

我们先讨论第二种解释. $\mathbb{C}$ 的任一非零元素 z 都可表示为极坐标形式 $z = r\mathrm{e}^{\mathrm{i}\theta}$, 其中 $r = |z|$, 而 $\theta = \arg(z)$. 如果 T 将 z 旋转一个角度 α, 则

$$T(z) = r\mathrm{e}^{\mathrm{i}(\theta+\alpha)} = \mathrm{e}^{\mathrm{i}\alpha}z$$

因此每个非零复数 z 都是属于特征值 $\lambda = \mathrm{e}^{\mathrm{i}\alpha} = \cos\alpha + \mathrm{i}\sin\alpha$ 的特征向量. 我们指出, 只有当 α 是 π 的整数倍时, $\mathrm{e}^{\mathrm{i}\alpha}$ 才是实数, 在其余情形, $\mathrm{e}^{\mathrm{i}\alpha}$ 都不是实数. 当 α 是 π 的偶数倍时, T 为恒等映射, 即 $T(z) = z$; 当 α 是 π 的奇数倍时, T 将每个点关于坐标原点作反射, 因此 $T(z) = -z$.

再来讨论第一种解释, 即将平面看作实线性空间 $\mathbb{R}^2$ 时的情形. 由于此时纯量只能取实数, 因此只有当旋转角 α 是 π 的整数倍时旋转变换 T 才有实特征值. 换言之, 若 α 不是 π 的整数倍, 则 T 没有实特征值, 从而也就没有特征向量. 由此可知, 特征值与特征向量的存在性可能依赖于 V 的纯量的选择. □

例 6 (由一个特征向量生成的子空间) 设 $T: S \to V$ 为一个线性变换, λ 为 T 的一个特征值. 令 x 为属于 λ 的一个特征向量, $L(x)$ 为由 x 生成的子空间, 即 $L(x)$ 为由 x 的全体纯量倍数所组成的子空间. 不难验证 T 将 $L(x)$ 映为它自身. 事实上, 若设 $y = cx$, 则有

$$T(y) = T(cx) = cT(x) = c(\lambda x) = \lambda(cx) = \lambda y.$$

若 $c \neq 0$, 则 $y \neq O$, 因此 $L(x)$ 中任一非零元素都是属于 λ 的特征向量.

设 U 为 S 的一个子空间, 若 T 将 U 中每一个元素都映为 U 中的元素, 则称 U 为 T 的一个**不变子空间** (invariant subspace), 或称 U 在 T 下**不变**. 由例 6 可知, 由一个特征向量生成的子空间在 T 下不变. □

下面两个例子需要用到微积分的知识.

例 7 (微分算子) 设 V 为由在某个开区间上任意阶可微的全体实函数 f 所组成的线性空间. 令 D 为将每个 f 映为其导数的线性变换, 即 $\mathrm{D}(f)=f'$, 则 D 的特征向量是 V 中对某个实数 λ 满足方程

$$f'=\lambda f$$

的非零实函数 f. 上述方程是一阶线性微分方程, 它的全部解由公式

$$f(x)=c\mathrm{e}^{\lambda x}$$

给出, 此处 c 为任意实常数. 从而 D 的特征向量是全体指数函数 $f(x)=c\mathrm{e}^{\lambda x}$, 其中 $c\neq 0$, 对应于特征向量 $f(x)=c\mathrm{e}^{\lambda x}$ 的特征值是 λ. 在与本例相类似的情形, 即当 V 是函数空间时, 我们常将特征向量叫作**特征函数** (eigenfunction). □

例 8 (积分算子) 设 v 是由在有限闭区间 $[a,b]$ 上连续的全体实函数所组成的线性空间. 对 $f\in V$, 由

$$g(x)=\int_a^x f(t)\mathrm{d}t,\quad a\leqslant x\leqslant b$$

定义函数 $g=T(f)$. 若 T 的特征函数存在, 则它必是对某个实数 λ 满足条件

$$\int_a^x f(t)\mathrm{d}t=\lambda f(x) \tag{6.3}$$

的非零函数 f. 若特征函数存在, 对方程 (6.3) 两边微分得 $f(x)=\lambda f'(x)$, 从而当 $\lambda\neq 0$ 时即得 $f(x)=c\mathrm{e}^{x/\lambda}$. 换言之, 只有形如 $f(x)=c\mathrm{e}^{x/\lambda}$ ($c\neq 0$ 且 $\lambda\neq 0$) 的函数才有可能是特征函数. 然而, 若在方程 (6.3) 中取 $x=a$, 则得

$$0=\lambda f(a)=\lambda c\mathrm{e}^{a/\lambda}.$$

由于 $\mathrm{e}^{a/\lambda}$ 不可能为零, 因此任意非零函数 f 都不可能满足方程 $T(f)=\lambda f$, 从而 T 没有特征函数, 也没有特征值. □

6.3 属于不同特征值的特征向量的线性无关性

下述定理给出特征向量的一个基本性质. 与上一节一样, S 表示线性空间 V 的一个子空间.

定理 6.2 设 $\lambda_1,\cdots,\lambda_k$ 为线性变换 $T:S\to V$ 的 k 个互不相同的特征值. 对 $1\leqslant i\leqslant k$, 令 u_i 为属于 λ_i 的特征向量, 则 $u_1,\cdots,u_k$ 线性无关.

证明 对特征向量的个数 k 作归纳法. 当 $k=1$ 时结论显然成立. 假定结论对任意 $k-1$ 个特征向量都成立. 令 $u_1,u_2,\cdots,u_k$ 分别为属于 k 个不同特征值 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量, 若有 k 个纯量 $c_1,c_2,\cdots,c_k$ 使

$$\sum_{i=1}^k c_iu_i=O. \tag{6.4}$$

将 T 作用于等式 (6.4) 的两边, 由 $T(u_i)=\lambda_iu_i$, 得

$$\sum_{i=1}^{k} c_i \lambda_i u_i = O. \tag{6.5}$$

将 λ_k 乘等式 (6.4) 的两边之后再与等式 (6.5) 相减得

$$\sum_{i=1}^{k-1} c_i (\lambda_i - \lambda_k) u_i = O.$$

由归纳假设, $u_1, \cdots, u_{k-1}$ 线性无关, 因此对 $i = 1, 2, \ldots, k-1$, 都有 $c_i(\lambda_i - \lambda_k) = 0$. 由于特征值互不相同, $(\lambda_i - \lambda_k) \neq 0$, 因此必有 $c_1 = c_2 = \cdots = c_{k-1} = 0$, 再代入等式 (6.4) 又得 $c_k = 0$, 因此特征向量 $u_1, u_2, \cdots, u_k$ 线性无关. □

我们指出, 如果允许将零元素也当作特征向量, 则定理 6.2 将不再成立. 这是我们不把零元素作为特征向量的第二个理由.

在有限维线性空间的情形, 定理 6.2 有下述重要推论.

定理 6.3 设 $\dim V = n$, 则任意一个线性变换 $T: V \to V$ 至多有 n 个不同的特征值. 如果 T 恰有 n 个不同的特征值, 则它们所对应的 n 个特征向量构成 V 的一组基, 并且 T 在这组基下的矩阵是以 T 的 n 个特征值作为主对角线元素的对角矩阵.

证明 如果存在 $n+1$ 个不同的特征值, 则由定理 6.2 可知 V 将包含 $n+1$ 个线性无关的元素, 这与 $\dim V = n$ 相矛盾, 因此 T 至多只能有 n 个不同的特征值. 第二个断言可由定理 6.1 和定理 6.2 推出. □

注 定理 6.3 告诉我们, 存在 n 个不同的特征值是 T 有一个对角矩阵表示的充分条件. 不过这个条件不是必要的. 存在这样的线性变换, 它只有少于 n 个不同的特征值, 但是它有一个对角矩阵表示. 恒等变换就是这样一个例子, 它的特征值都是 1, 但是恒等变换在任意一组基下的矩阵都是单位矩阵. 定理 6.1 给出了 T 有一个对角矩阵表示的充分必要条件.

6.4 习 题

1. (a) 证明: 若 T 有一个特征值 λ, 则 aT 有特征值 $a\lambda$.
   (b) 证明: 若 x 既是 T_1 的特征向量, 又是 T_2 的特征向量, 则对所有的纯量 a 和 b, x 都是 $aT_1 + bT_2$ 的特征向量. 它们对应的特征值之间有什么关系?
2. 设 $T: V \to V$ 有一个属于特征值 λ 的特征向量 x, 证明 x 也是 T^2 的属于特征值 λ^2 的特征向量. 一般地, 对任意正整数 k, x 也是 T^k 的属于特征值 λ^k 的特征向量. 再利用习题 1 的结果证明: 对任意多项式 P, x 也是 $P(T)$ 的属于特征值 $P(\lambda)$ 的特征向量.
3. 将平面看作实线性空间 $\mathbb{R}^2$, 令 T 为 $\mathbb{R}^2$ 的绕原点旋转 $\pi/2$ 角度的旋转变换. 证明: 尽管 T 没有特征向量, 然而每一个非零向量都是 T^2 的特征向量.
4. 设 $T: V \to V$ 有这样的性质: T^2 有一个非负特征值 λ^2. 证明: λ 与 $-\lambda$ 中至少有一个是 T 的特征值. [提示: $T^2 - \lambda^2 I = (T + \lambda I)(T - \lambda I)$.]
5. 设 x 与 y 分别为线性变换 T 的属于特征值 λ 与 μ 的特征向量, 其中 $\lambda \neq \mu$. 证明: 若 $ax + by$ 为 T 的特征向量, 则必有 $a = 0$ 或 $b = 0$.

6. 设 $T: S \to V$ 是一个线性变换, 它使得 S 中任意一个非零元素都是它的特征向量. 证明: 存在纯量 c, 使对所有的 $x \in S$ 都有 $T(x) = cx$. 换言之, 具有上述性质的线性变换一定能够表示成恒等变换与某个纯量之积 (提示: 应用习题 5 的结论).
7. 令 V 为由全体次数 $\leqslant n$ 的实系数多项式 $p(x)$ 以及零多项式所构成的线性空间. 对 $p \in V$, 定义 $q = T(p)$ 如下: 对所有实数 t 都有 $q(t) = p(t+1)$. 证明 T 的特征值只能是 1, 并求出属于此特征值的全部特征函数.

   习题 8~11 要用到微积分的知识.
8. 设 V 是由区间 $(0,1)$ 上全体可微实函数所构成的线性空间. 对 $f \in V$, 定义 $g = T(f)$ 如下: 对任意 $t \in (0,1)$, 令 $g(t) = tf'(t)$. 证明任意实数 λ 都是 T 的特征值并确定属于 λ 的特征函数.
9. 设 V 是由区间 $(-\infty, +\infty)$ 上全体连续函数所构成的线性空间, 则对任意 $f \in V$ 及任意实数 x, 积分 $\int_{-\infty}^{x} f(t)\mathrm{d}t$ 都存在. 对 $f \in V$, 依如下方式定义 $g = T(f): g(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$. 证明每一个正实数 λ 都是 T 的特征值并求出属于特征值 λ 的特征函数.
10. 设 V 是由区间 $(-\infty, +\infty)$ 上全体连续函数所构成的线性空间, 则对任意 $f \in V$ 及任意实数 x, 积分 $\int_{-\infty}^{x} tf(t)\mathrm{d}t$ 都存在. 对 $f \in V$, 依如下方式定义 $g = T(f): g(x) = \int_{-\infty}^{x} tf(t)\mathrm{d}t$. 证明每一个负实数 λ 都是 T 的特征值并求出属于特征值 λ 的特征函数.
11. 设 V 为由全体实收敛序列 $\{x_n\}$ 所构成的线性空间. 依如下方式定义 $T: V \to V$: 若 $x = \{x_n\}$ 是一个以 a 为极限的收敛序列, 则令 $T(x) = \{y_n\}$, 此处 $y_n = a - x_n, n \geqslant 1$. 证明 $\lambda = 0$ 与 $\lambda = -1$ 是 T 所仅有的两个特征值并求出分别属于这两个特征值的特征向量.

6.5 有限维线性空间

在 6.2 节例 5 中我们看到, 对一个给定的线性算子, 特征值的存在性与纯量域有关. 下述定理指出, 如果取复数作为纯量并且线性空间 V 是有限维的, 即如果 V 是有限维复线性空间, 则特征值必定存在.

定理 6.4 设 V 是一个有限维复线性空间, 则任一线性变换 $T: V \to V$ 必定有特征值存在.

证明 令 $n = \dim V$, 任意取定 V 中的一个非零元素 x, 考虑 V 中下述 $n+1$ 个元素

$$x, T(x), T^2(x), \cdots, T^n(x).$$

由于 $\dim V = n$, 因此上述 $n+1$ 个元素线性相关. 从而存在 $n+1$ 个不全为零的复数 $c_0, c_1, \cdots, c_n$ 使得

$$\sum_{k=0}^{n} c_k T^k(x) = O.$$

上式也可表为 $F(x) = O$ 的形式, 其中 F 为下述线性算子:

$$F = \sum_{k=0}^{n} c_k T^k.$$

以各个复数 c_k 为系数作如下的复值多项式:

$$f(z)=\sum_{k=0}^{n}c_kz^k.$$

若 c_m 是各系数 $c_0,c_1,\cdots,c_n$ 中的最后一个非零系数, 则 $f(z)$ 是一个 m 次多项式, 此处 $m\leqslant n$. 将 $f(z)$ 分解为一次因式乘积的形式

$$f(z)=c_m(z-\alpha_1)\cdots(z-\alpha_m),$$

其中 $\alpha_1,\cdots,\alpha_m$ 是多项式 $f(z)$ 的 m 个复根, 则相应的线性算子 F 也可分解成如下形式:

$$F=c_m(T-\alpha_1I)\cdots(T-\alpha_mI),$$

其中 I 是恒等变换且 $c_m\neq 0$. 将 F 作用在非零元素 x 上, 我们得

$$F(x)=c_m(T-\alpha_1I)\cdots(T-\alpha_mI)(x)=O.$$

令 $x_1=(T-\alpha_mI)(x)$, 若 $x_1=O$, 则得 $T(x)=\alpha_mx$, 因此 α_m 是一个特征值, x 是属于 α_m 的一个特征向量. 若 $x_1\neq O$, 则依次定义元素 $x_2,x_3,\cdots,x_m$ 如下:

$$x_2=(T-\alpha_{m-1}I)(x_1),x_3=(T-\alpha_{m-2}I)(x_2),\cdots,x_m=(T-\alpha_1I)(x_{m-1}).$$

由于 $F(x)=O$ 且 $c_m\neq 0$, 因此 $x_2,x_3,\cdots,x_m$ 中至少有一个为零元素. 令 x_k 为 $x_2,x_3,\cdots,x_m$ 中最后一个非零元素, 则 $x_{k+1}=O$, 因此 $T(x_k)=\alpha_{m-k}x_k$, 于是 x_k 便是属于特征值 α_{m-k} 的一个特征向量. □

6.6 三角化定理

由定理 6.1 知, 若 T 为作用在 n 维线性空间 V 上的一个线性变换, 则 T 可对角化的充分必要条件是 T 有 n 个线性无关的特征向量. 于是, 若 T 没有 n 个线性无关的特征向量, 则 T 便不可能对角化. 不过, 如果 V 是 n 维复线性空间, 则由下述定理可知, 我们总可以将 T **三角化** (Triangularized), 即使 T 的线性无关的特征向量的总数少于 n 也不例外.

定理 6.5 (三角化定理) 设 V 是一个 n 维复线性空间, $T:V\to V$ 为任一线性变换, 则总存在 V 的一组基, 使得 T 在这组基下的矩阵为上三角形矩阵, 并且这个上三角形矩阵以 T 的 n 个特征值作为它的主对角线元素.

证明 对 V 的维数 n 作归纳法. 当 $n=1$ 时, 由于任意 1×1 矩阵都是上三角形的, 并且它仅有的一个元素就是它的特征值, 因此结论显然成立. 假定对所有作用在 $n-1$ 维复线性空间上的线性变换结论都成立, 我们要证明对所有作用在 n 维复线性空间上的线性变换结论也都成立.

由定理 6.4, 复线性空间 V 上的线性变换 T 必定存在特征值, 故不妨设 λ_1 为 T 的一个特征值, 再令 u_1 为属于 λ_1 的一个特征向量, 则有

$$T(u_1)=\lambda_1u_1 \text{ 且 } u_1\neq O.$$

由定理 3.7(a) 知, u_1 可扩充为 V 的一组基, 设这组基为

$$U = (u_1, u_2, \cdots, u_n).$$

令 $\boldsymbol{A} = (a_{ij})$ 为 T 在这组基下的矩阵, 则可设

$$T(u_k) = \sum_{i=1}^{n} a_{ik} u_i \quad k = 1, 2, \cdots, n. \tag{6.6}$$

令 $V^* = L(u_2, \cdots, u_n)$ 为由 $n-1$ 个元素 $u_2, \cdots, u_n$ 生成的线性空间. 由于 $u_2, \cdots, u_n$ 线性无关, 因此 $\dim V^* = n-1$. 为应用归纳假设, 我们定义一个新的线性变换 $T^*: V^* \to V^*$, T^* 在 V^* 的基元素上的作用由下式给出:

$$T^*(u_k) = \sum_{i=2}^{n} a_{ik} u_i, \quad k = 2, \cdots, n. \tag{6.7}$$

注意到和式 (6.7) 是将和式 (6.6) 中的第一项去掉后所得, 只有元素 $u_2, \cdots, u_n$ 仍然保留, 因此 T^* 将 V^* 映入 V^*, 从而易知 T^* 是 $n-1$ 维复线性空间 V^* 上的一个线性变换, 由归纳假设可知, 存在 V^* 的一组基, 设为

$$W^* = (w_2, \cdots, w_n),$$

使得 T^* 在这组基下的矩阵是一个上三角形矩阵, 将它记作 (b_{ij}), 其中 $2 \leqslant i, j \leqslant n$. 于是我们得

$$\begin{aligned} T^*(w_2) &= b_{22} w_2, \\ T^*(w_3) &= b_{23} w_2 + b_{33} w_3, \\ &\vdots \\ T^*(w_n) &= b_{2n} w_2 + b_{3n} w_3 + \cdots + b_{nn} w_n, \end{aligned}$$

由归纳假设, (b_{ij}) 的主对角线元素 $b_{22}, b_{33}, \cdots, b_{nn}$ 都是 T^* 的特征值.

令 $W = (u_1, w_2, \cdots, w_n)$, 我们要证明 W 是 V 的一组基, T 在 W 下的矩阵为上三角形矩阵. 由于 W 包含 n 个 V 中元素, 因此为证明 W 是 V 的一组基, 只需证明 W 生成 V 即可. 若 $x \in V$, 则 x 是 $u_1, u_2, \cdots, u_n$ 的一个线性组合. 又因 $u_2, \cdots, u_n$ 都是 V^* 中的元素, 因此它们都可表示为 $w_2, \cdots, w_n$ 的线性组合, 从而 x 可表示为 $u_1, w_2, \cdots, w_n$ 的线性组合, 亦即 W 生成线性空间 V.

为求 T 在基 W 之下的矩阵, 我们考察基元素在 T 作用下的象. 对第一个基元素 u_1, 我们已知有

$$T(u_1) = \lambda_1 u_1.$$

为求 $w_2, \cdots, w_n$ 在 T 作用下的象, 我们来比较 T 与 T^* 在 V^* 的一个典型元素上的作用. 若 $x \in V^*$, 则存在复数 $c_2, c_3, \cdots, c_n$, 使得

$$x = \sum_{k=2}^{n} c_k u_k,$$

从而得

$$T(x) = \sum_{k=2}^{n} c_k T(u_k),$$

$$T^*(x) = \sum_{k=2}^{n} c_k T^*(u_k).$$

于是上述二式之差为

$$T(x) - T^*(x) = \sum_{k=2}^{n} c_k \{T(u_k) - T^*(u_k)\}.$$

由 (6.6) 与 (6.7) 二式得 $T(u_k) - T^*(u_k) = a_{1k}u_1$, 因此

$$T(x) - T^*(x) = \left(\sum_{k=2}^{n} c_k a_{1k}\right) u_1 = c(x)u_1,$$

此处

$$c(x) = \sum_{k=2}^{n} c_k a_{1k}$$

是某个依赖于 x 的复数. 换言之, 我们有

$$T(x) = c(x)u_1 + T^*(x). \tag{6.8}$$

依次取 $x = w_2, \cdots, w_n$, 反复应用上述关系式, 我们得

$$\begin{aligned}
T(w_2) &= c(w_2)u_1 + T^*(w_2) = c(w_2)u_1 + b_{22}w_2, \\
T(w_3) &= c(w_3)u_1 + T^*(w_3) = c(w_3)u_1 + b_{23}w_2 + b_{33}w_3, \\
&\vdots \\
T(w_n) &= c(w_n)u_1 + T^*(w_n) = c(w_n)u_1 + b_{2n}w_2 + b_{3n}w_3 + \cdots + b_{nn}w_n.
\end{aligned}$$

因此 T 在基 W 下的矩阵是一个以 $\lambda_1, b_{22}, \cdots, b_{nn}$ 为主对角线元素的上三角形矩阵.

余下还需证明此上三角形矩阵的各个主对角线元素都是 T 的特征值. 前面已知 λ_1 是 T 的特征值, 又由归纳假设, $b_{22}, \cdots, b_{nn}$ 都是 T^* 的特征值, 因此在 V^* 中存在特征向量 $v_2, v_3, \cdots, v_n$ 使得 $T^*(v_k) = b_{kk}v_k, \quad k = 2, 3, \cdots, n$. 由等式 (6.8) 得

$$T(v_k) = c(v_k)u_1 + T^*(v_k) = c(v_k)u_1 + b_{kk}v_k.$$

若 $c(v_k) = 0$, 则 $T(v_k) = b_{kk}v_k$, 从而 v_k 是 T 的一个属于特征值 b_{kk} 的特征向量. 而若 $c(v_k) \neq 0$, 则 v_k 不是 T 的特征向量, 不过我们可以通过适当选取复数 λ_k 使得

$$v_k' = u_1 + \lambda_k v_k$$

为 T 的一个特征向量. 由于 u_1 与 v_k 线性无关, 因此对任意选取的 λ_k, 元素 v_k' 都不可能是零元素. 为选取使 v_k' 为特征向量的 λ_k, 将 T 作用在 v_k' 上得

$$\begin{aligned}T(v_k')&=T(u_1)+\lambda_kT(v_k)=\lambda_1u_1+\lambda_k\{c(v_k)u_1+b_{kk}v_k\}\\&=\{\lambda_1+\lambda_kc(v_k)\}u_1+b_{kk}\{v_k'-u_1\}\\&=b_{kk}v_k'+\{\lambda_1+\lambda_kc(v_k)-b_{kk}\}u_1.\end{aligned}$$

当 $c(v_k)\neq 0$ 时, 存在唯一的一个 λ_k, 即 $\lambda_k=(b_{kk}-\lambda_1)/c(v_k)$, 使得 $\{\lambda_1+\lambda_kc(v_k)-b_{kk}\}=0$, 对这个 λ_k, 我们有

$$T(v_k')=b_{kk}v_k';$$

因此 v_k' 是一个属于特征值 b_{kk} 的特征向量. 因此, T 在基 W 下的矩阵的主对角线元素 $\lambda_1, b_{22}, \cdots, b_{nn}$ 都是 T 的特征值, 由此完成了三角化定理的证明. □

6.7 特征多项式

我们接下来研究当线性变换 $T:V\to V$ 的特征值存在时如何实际上求出它们的问题. 我们寻找使方程 $T(x)=\lambda x$ 有非零解 x 的纯量 λ. 方程 $T(x)=\lambda x$ 可写成如下形式:

$$(\lambda I-T)(x)=O, \tag{6.9}$$

此处 I 是恒等变换. 于是当且仅当方程 (6.9) 有非零解 x 时, λ 是 T 的一个特征值, 由定理 4.10, 此时 $\lambda I-T$ 是不可逆的.

当 V 是有限维线性空间时, 我们可以借助于行列式来求特征值. 设 $\boldsymbol{A}$ 是线性变换 T 的一个矩阵表示, 则 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 是线性变换 $\lambda I-T$ 的一个矩阵表示, 而且线性变换 $\lambda I-T$ 可逆的充分必要条件是 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 为非奇异矩阵. 于是, λ 是 T 的特征值的充分必要条件是 $\lambda\boldsymbol{I}-\boldsymbol{A}$ 为奇异矩阵. 由定理 5.17, λ 是 T 的特征值的充分必要条件是 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$. 换言之, T 的每一个特征值 λ 都满足方程

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0. \tag{6.10}$$

反之, 任一满足条件 (6.10) 的纯量 λ 都是 T 的特征值. 这建议我们将行列式 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ 当作 λ 的一个函数来进行研究.

定理 6.6 *设 $\boldsymbol{A}=(a_{ij})$ 是一个 $n\times n$ 矩阵, $\boldsymbol{I}$ 为 $n\times n$ 单位矩阵, 则由*

$$f(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})$$

定义的函数 f 是 λ 的一个 n 次多项式, 而且这个多项式的最高次项是 λ^n, 其常数项是 $f(0)=\det(-\boldsymbol{A})=(-1)^n\det\boldsymbol{A}$.

证明 由 f 的定义即得 $f(0)=\det(-\boldsymbol{A})$. 在 $n\leqslant 3$ 的情形我们证明 f 是 λ 的 n 次多项式. 对一般的 n 可用归纳法给出证明, 我们将它作为习题留给读者 (见 6.10 节习题 9).

当 $n=1$ 时, $f(\lambda)=\lambda-a_{11}$ 是 λ 的一次多项式, 当 $n=2$ 时, 我们有

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-a_{11} & -a_{12}\\ -a_{21} & \lambda-a_{22}\end{vmatrix}=(\lambda-a_{11})(\lambda-a_{22})-a_{12}a_{21}$$
$$=\lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-a_{12}a_{21}),$$

这是 λ 的一个二次多项式. 当 $n=3$ 时, 我们有

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-a_{11} & -a_{12} & -a_{13}\\ -a_{21} & \lambda-a_{22} & -a_{23}\\ -a_{31} & -a_{32} & \lambda-a_{33}\end{vmatrix}$$
$$=(\lambda-a_{11})\begin{vmatrix}\lambda-a_{22} & -a_{23}\\ -a_{32} & \lambda-a_{33}\end{vmatrix}+a_{12}\begin{vmatrix}-a_{21} & -a_{23}\\ -a_{31} & \lambda-a_{33}\end{vmatrix}$$
$$-a_{13}\begin{vmatrix}-a_{21} & \lambda-a_{22}\\ -a_{31} & -a_{32}\end{vmatrix},$$

上式后两项都是 λ 的一次多项式, 第一项是 λ 的三次多项式, 其最高次项为 λ^3, 故这是 λ 的一个三次多项式. □

定义　设 $\boldsymbol{A}$ 是一个 $n\times n$ 矩阵, 行列式

$$f(\lambda)=\det(\lambda \boldsymbol{I}-\boldsymbol{A})$$

叫作 $\boldsymbol{A}$ 的**特征多项式** (characteristic polynomial).

$\boldsymbol{A}$ 的特征多项式的根都是复数, 其中某一些可能是实数, 如果令 F 表示实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$, 则可得下述定理.

定理 6.7　设 V 为 n 维线性空间, 其纯量域为 F, 设 $T:V\to V$ 为线性变换, 设 $\boldsymbol{A}$ 为 T 的某个矩阵表示, 则 T 的特征值的集合由 $\boldsymbol{A}$ 的特征多项式的属于 F 的根组成.

证明　由定理 6.6 之前的讨论知, T 的每一个特征值都满足方程 $\det(\lambda \boldsymbol{I}-\boldsymbol{A})=0$, 而且 $\boldsymbol{A}$ 的特征多项式的任意一个属于 F 的根都是 T 的特征值. 故得定理. □

矩阵 $\boldsymbol{A}$ 依赖于 V 的基的选择, 而 T 的特征值的定义与基无关, 因此 $\boldsymbol{A}$ 的特征多项式的根的集合必须与基的选择无关. 不仅如此, 我们还将在 6.11 节中证明, 特征多项式本身也和基的选择无关.

我们下面讨论在有限维的情形下特征值与特征向量的实际求法.

6.8　有限维情形下特征值与特征向量的计算

在有限维情形, 线性变换 T 的特征值和特征向量也叫作 T 的任何一个矩阵表示的特征值和特征向量. 于是方阵 $\boldsymbol{A}$ 的特征值就是其特征多项式 $f(\lambda)=\det(\lambda \boldsymbol{I}-\boldsymbol{A})$ 的属于纯量域 F 的根. 属于特征值 λ 的特征向量就是那些满足条件

$$\boldsymbol{A}\boldsymbol{X}=\lambda\boldsymbol{X}\quad 即\quad (\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{O}$$

的非零列向量 $\boldsymbol{X}=(x_1,\cdots,x_n)^{\mathrm{T}}$. 上式是关于 $x_1,\cdots,x_n$ 的由 n 个线性方程组成的方程组. 只要知道 λ, 我们就能通过解这个线性方程组来求出特征向量. 我们给出三个不同类型的例子来说明特征向量的求法.

例 1 (所有特征值都不同的矩阵) 矩阵

$$\boldsymbol{A}=\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

的特征多项式为

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda-2 & -1 & -1 \\ -2 & \lambda-3 & -4 \\ 1 & 1 & \lambda+2 \end{vmatrix}=(\lambda-1)(\lambda+1)(\lambda-3),$$

因此 $\boldsymbol{A}$ 有 3 个不同的特征值: $1,-1,3$. 为求属于 $\lambda=1$ 的特征向量, 我们解线性方程组 $\boldsymbol{AX}=\boldsymbol{X}$, 即

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

它给出由 3 个方程组成的线性方程组

$$\begin{cases} 2x_1+x_2+x_3=x_1 \\ 2x_1+3x_2+4x_3=x_2 \\ -x_1-x_2-2x_3=x_3 \end{cases}$$

可将它改写成如下形式:

$$\begin{cases} x_1+x_2+x_3=0 \\ 2x_1+2x_2+4x_3=0 \\ -x_1-x_2-3x_3=0 \end{cases}$$

将第一个方程和第三个方程相加便得 $x_3=0$, 于是所有上述三个方程可化为 $x_1+x_2=0$, 因此属于 $\lambda=1$ 的特征向量为 $\boldsymbol{X}=t(1,-1,0)$, 此处 t 为任意非零纯量.

同法可得属于特征值 $\lambda=-1$ 的特征向量为 $\boldsymbol{X}=t(0,1,-1)$, 属于 $\lambda=3$ 的特征向量为 $\boldsymbol{X}=t(2,3,-1)$, 此处 t 为任意不等于零的纯量. 由于特征值互不相同, 因此相应的特征向量线性无关. 所得结果总结成下表, 在表的第三列中我们列出了特征空间 $E(\lambda)$ 的维数.

特征值 λ	特征向量	$E(\lambda)$ 的维数
1	$t(1,-1,0),t\neq 0$	1
-1	$t(0,1,-1),t\neq 0$	1
3	$t(2,3,-1),t\neq 0$	1

□

例 2 (有 2 重特征值的矩阵) 矩阵

$$A=\begin{bmatrix}2 & -1 & 1\\0 & 3 & -1\\2 & 1 & 3\end{bmatrix}$$

的特征多项式为

$$\det(\lambda I-A)=\begin{vmatrix}\lambda-2 & 1 & -1\\0 & \lambda-3 & 1\\-2 & -1 & \lambda-3\end{vmatrix}=(\lambda-2)(\lambda-2)(\lambda-4),$$

因此特征值为 2, 2, 4(我们将特征值 2 重复列出, 这是为了强调它是特征多项式的 2 重根). 为求属于 $\lambda=2$ 的特征向量, 我们解方程组 $AX=2X$, 它可改写为如下形式:

$$\begin{cases}-x_2+x_3=0\\x_2-x_3=0\\2x_1+x_2+x_3=0.\end{cases}$$

该方程组的通解为 $x_2=x_3=-x_1$, 因此属于 $\lambda=2$ 的特征向量为 $t(-1,1,1)$, 其中 $t\neq 0$. 同法可得属于特征值 $\lambda=4$ 的特征向量为 $t(1,-1,1), t\neq 0$. 将上述结果总结成下表:

特征值	特征向量	$E(\lambda)$的维数
2, 2	$t(-1,1,1), t\neq 0$	1
4	$t(1,-1,1), t\neq 0$	1

□

例 3 (另一个有 2 重特征值的矩阵) 矩阵

$$A=\begin{bmatrix}2 & 1 & 1\\2 & 3 & 2\\3 & 3 & 4\end{bmatrix}$$

的特征多项式为 $(\lambda-1)(\lambda-1)(\lambda-7)$. 当 $\lambda=7$ 时, 方程组 $AX=7X$ 可改写为

$$\begin{cases}5x_1-x_2-x_3=0\\-2x_1+4x_2-2x_3=0\\-3x_1-3x_2+3x_3=0,\end{cases}$$

其解为 $x_2=2x_1, x_3=3x_1$, 因此属于特征值 $\lambda=7$ 的特征向量为 $t(1,2,3), t\neq 0$. 对特征值 $\lambda=1$, 方程组 $AX=X$ 可化为由 3 个相同的方程 $x_1+x_2+x_3=0$ 组成的齐次线性方程组. 为求其解, 可取 $x_1=a, x_2=b$, 其中 a 与 b 可任意取, 而取 $x_3=-a-b$. 因此属于 $\lambda=1$ 的每一个特征向量都可表示为如下形式:

$$(a,b,-a-b)=a(1,0,-1)+b(0,1,-1),$$

其中 $a \neq 0$ 或 $b \neq 0$. 这表示线性无关向量组 $(1,0,-1)$ 与 $(0,1,-1)$ 组成 $E(1)$ 的一组基. 因此当 $\lambda = 1$ 时, $\dim E(\lambda) = 2$. 上述结果可整理成下表:

特征值	特征向量	$E(\lambda)$的维数
7	$t(1,2,3), t \neq 0$	1
1, 1	$a(1,0,-1)+b(0,1,-1), a,b$ 不同时为 0	2

注　在本例中, 虽然只有两个不同的特征值, 但却有 3 个线性无关的特征向量. □

6.9 特征多项式根的积与和

设 $f(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A})$ 为 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 的特征多项式, 令 $\lambda_1, \cdots, \lambda_n$ 表示 $f(\lambda)$ 的 n 个根, 每个根出现的次数等于它的重数 (2 重根重复 2 次, 3 重根重复 3 次等). 则可将 $f(\lambda)$ 表示为一次因式的乘积

$$f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

我们也可以将多项式 $f(\lambda)$ 按 λ 的降幂排列:

$$f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

将此式和前面表为一次因式乘积的分解式相比较, 可知常数项 c_0 及 λ^{n-1} 的系数分别由公式

$$c_0 = (-1)^n \lambda_1 \cdots \lambda_n, \quad c_{n-1} = -(\lambda_1 + \cdots + \lambda_n)$$

给出, 由于我们也有 $c_0 = f(0) = (-1)^n \det \boldsymbol{A}$, 因此得

$$\lambda_1 \cdots \lambda_n = \det \boldsymbol{A}$$

换言之, 我们有下述结论:

$\boldsymbol{A}$ 的特征多项式的 n 个根的乘积等于 $\boldsymbol{A}$ 的行列式.

重根 (特征值) 在这个乘积中出现的次数必须和重因式在 $f(\lambda)$ 的一次因式乘积的分解式中出现的次数相同. 例如, 在 6.8 节例 2 中, 矩阵 $\boldsymbol{A}$ 的特征值为 2, 2, 4, 因此 $\boldsymbol{A}$ 的全体特征值 (考虑重数) 的乘积是 $2 \times 2 \times 4 = 16$, 与 $\det \boldsymbol{A} = 16$ 相一致.

$f(\lambda)$ 的所有根的和叫做 $\boldsymbol{A}$ 的**迹** (trace), 记作 $\mathrm{tr}\boldsymbol{A}$, 因此由定义得

$$\mathrm{tr}\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i,$$

这里, 重根依它的重数重复出现. λ^{n-1} 的系数由 $c_{n-1} = -\mathrm{tr}\boldsymbol{A}$ 给出. 我们也可由 $f(\lambda)$ 的行列式形式求出这个系数为 (见 6.10 节习题 12):

$$c_{n-1} = -(a_{11} + \cdots + a_{nn}).$$

由此得

$$\mathrm{tr}\boldsymbol{A} = \sum_{i=1}^{n} a_{ii},$$

即矩阵的迹等于其主对角线上的元素之和.

由于一个矩阵主对角元素之和很容易计算出, 因此它常被用来验证关于特征值的计算. 关于迹的进一步性质放在下面的一组习题中讨论.

6.10 习　　题

求出习题 1~3 中各矩阵的特征值与特征向量, 并对每个特征值 λ 求特征空间 $E(\lambda)$ 的维数.

1. (a) $\begin{bmatrix}1&0\\0&1\end{bmatrix}$.　(b) $\begin{bmatrix}1&1\\0&1\end{bmatrix}$.　(c) $\begin{bmatrix}1&0\\1&1\end{bmatrix}$.　(d) $\begin{bmatrix}1&1\\1&1\end{bmatrix}$.

2. $\begin{bmatrix}1&a\\b&1\end{bmatrix}, a>0, b>0$.

3. $\begin{bmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{bmatrix}$.

4. 矩阵

$$\boldsymbol{P}_1=\begin{bmatrix}0&1\\1&0\end{bmatrix}, \boldsymbol{P}_2=\begin{bmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{bmatrix}, \boldsymbol{P}_3=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$$

在关于电子自旋的量子力学理论中用到, 为纪念物理学家 Wolfgang Pauli(1900 − 1958) 而将它们叫作 Pauli 自旋矩阵. 验证这些矩阵的特征值都是 1 与 −1. 然后确定以 1 与 −1 为特征值的全体 2×2 复矩阵.

5. 确定所有这样的 2×2 矩阵, 其特征值为:

   (a) 一对不同的实数;　(b) 一对相同的实数;　(c) 一对共轭复数.

6. 求 a,b,c,d,e,f 使得向量 $(1,1,1)$, $(1,0,-1)$ 与 $(1,-1,0)$ 为矩阵

$$\begin{bmatrix}1&1&1\\a&b&c\\d&e&f\end{bmatrix}$$

的特征向量.

7. 求下列各矩阵的特征值与特征向量, 并对每个特征值 λ, 求特征子空间 $E(\lambda)$ 的维数.

   (a) $\begin{bmatrix}1&0&0\\-3&1&0\\4&-7&1\end{bmatrix}$　(b) $\begin{bmatrix}2&1&3\\1&2&3\\3&3&20\end{bmatrix}$　(c) $\begin{bmatrix}5&-6&-6\\-1&4&2\\3&-6&-4\end{bmatrix}$

8. 求下列各矩阵的特征值.

   (a) $\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix}$.　(b) $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix}$.　(c) $\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$.

   (d) $\begin{bmatrix}0&-\mathrm{i}&0&0\\\mathrm{i}&0&0&0\\0&0&0&-\mathrm{i}\\0&0&\mathrm{i}&0\end{bmatrix}$.　(e) $\begin{bmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}$.

   为纪念英国物理学家 Paul M. Dirac (1902 1984), 将这 5 个矩阵叫作 **Dirac 矩阵**, 它们在求解量子力学中相对论性波动方程时有用.

9. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为 $n\times n$ 矩阵, 其中 $\boldsymbol{B}$ 为对角矩阵, (用归纳法) 证明行列式 $f(\lambda)=\det(\lambda\boldsymbol{B}-\boldsymbol{A})$ 是关于 λ 的多项式, 使得 $f(0)=(-1)^n\det\boldsymbol{A}$ 并且 λ^n 的系数等于 $\boldsymbol{B}$ 的主对角线元素的乘积.

10. 证明方阵 $\boldsymbol{A}$ 与其转置矩阵 $\boldsymbol{A}'$ 有相同的特征多项式.
11. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为 $n\times n$ 矩阵, $\boldsymbol{A}$ 为非奇异矩阵, 证明 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征值.
    注: 即使 $\boldsymbol{A}$ 是奇异矩阵, 我们也可以证明 $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 有相同的特征多项式. 不过我们不要求读者证明这一结论.
12. 设 $\boldsymbol{A}$ 是 $n\times n$ 矩阵, $\boldsymbol{A}$ 的特征多项式为 $f(\lambda)$. (由归纳法) 证明在 $f(\lambda)$ 中 λ^{n-1} 的系数是 $-\mathrm{tr}\boldsymbol{A}$.
13. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为 $n\times n$ 矩阵, 并且有 $\det\boldsymbol{A}=\det\boldsymbol{B}$ 及 $\mathrm{tr}\boldsymbol{A}=\mathrm{tr}\boldsymbol{B}$, 证明: 当 $n=2$ 时, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 但当 $n>2$ 时这个结论并不总是成立.
14. 证明关于矩阵迹的下列各个结论.
    (a) $\mathrm{tr}(\boldsymbol{A}+\boldsymbol{B})=\mathrm{tr}\boldsymbol{A}+\mathrm{tr}\boldsymbol{B}$.
    (b) $\mathrm{tr}(c\boldsymbol{A})=c\,\mathrm{tr}\boldsymbol{A}$.
    (c) $\mathrm{tr}(\boldsymbol{AB})=\mathrm{tr}(\boldsymbol{BA})$.
    (d) $\mathrm{tr}\boldsymbol{A}^{\mathrm{T}}=\mathrm{tr}\boldsymbol{A}$.

6.11 表示同一个线性变换的矩阵 · 相似矩阵

我们将在本节中证明, 表示同一个线性变换的不同矩阵有相同的特征多项式. 为此, 我们将更深入地探究表示同一个线性变换的矩阵之间的关系.

先回顾关于线性变换矩阵表示的定义. 设 $T:V\to W$ 是从 n 维空间 V 到 m 维空间 W 的一个线性映射. 令 $(v_1,\cdots,v_n)$ 与 $(w_1,\cdots,w_m)$ 分别为 V 与 W 的有序基. 令 $\boldsymbol{A}$ 为一个 $m\times n$ 矩阵, 它分别以 $T(v_1),\cdots,T(v_n)$ 在基 $(w_1,\cdots,w_m)$ 表示下的坐标向量作为第 1 列, 第 2 列, $\cdots$, 第 n 列, 则称 A 为 T 在基 $(v_1,\cdots,v_n)$ 与 $(w_1,\cdots,w_m)$ 下的矩阵表示或矩阵. 若选取的基不同, 则所得到的矩阵表示通常也不同.

下面考虑 V 与 W 是同一个线性空间并且它们取同一组有序基 $(v_1,\cdots,v_n)$ 的情形. 若 $\boldsymbol{A}=(a_{ij})$ 是 T 在这组基下的矩阵, 则我们有

$$T(v_k)=\sum_{i=1}^{n}a_{ik}v_i,\quad k=1,2,\cdots,n. \tag{6.11}$$

若我们对同一空间 V 选取另一组有序基 $(u_1,\cdots,u_n)$ 并设 $\boldsymbol{B}=(b_{ij})$ 为 T 在这组新基下的矩阵, 我们将证明 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间存在一个简单的矩阵关系

$$\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{AC}, \tag{6.12}$$

其中 $\boldsymbol{C}$ 为某个非奇异矩阵.

为确定矩阵 $\boldsymbol{C}$, 我们注意到, 由 $\boldsymbol{B}$ 的定义得

$$T(u_j)=\sum_{k=1}^{n}b_{kj}u_k,\quad j=1,2,\cdots,n. \tag{6.13}$$

由于各个 u_j 都是由 $v_1,\cdots,v_n$ 生成的空间中的元素, 故都可以表示为 $v_1,\cdots,v_n$ 的线性组合. 设为

$$u_j=\sum_{k=1}^{n}c_{kj}v_k,\quad j=1,2,\cdots,n. \tag{6.14}$$

其中 c_{kj} 为纯量. 由这 n^2 个纯量组成的 $n\times n$ 矩阵 $\boldsymbol{C}=(c_{kj})$ 叫作**基变换矩阵**或**转移矩阵** (transition matrix), 它将基 $v_1,\cdots,v_n$ 变换为基 $u_1,\cdots,u_n$. 矩阵 $\boldsymbol{C}$ 表示了将一组基映到另一组基上的一个线性变换, 因此 $\boldsymbol{C}$ 是一个非奇异矩阵, 从而 $\boldsymbol{C}^{-1}$ 存在. 我们要证明矩阵等式 (6.12) 对这个 $\boldsymbol{C}$ 成立.

关系式 (6.14) 建立了两组基之间的联系, 它可以改写为下述简单的矩阵等式:

$$\boldsymbol{U}=\boldsymbol{V}\boldsymbol{C}, \tag{6.15}$$

其中 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 是两个 $1\times n$ 行矩阵, 它们的分量不是纯量而是两组相应的基元素:

$$\boldsymbol{U}=[u_1,\cdots,u_n],\quad \boldsymbol{V}=[v_1,\cdots,v_n].$$

矩阵等式 (6.15) 等价于 (6.14) 中的 n 个等式. 在等式 (6.15) 两边右乘 $\boldsymbol{C}^{-1}$ 得

$$\boldsymbol{V}=\boldsymbol{U}\boldsymbol{C}^{-1}. \tag{6.16}$$

(6.11) 式也可以表示成如下矩阵等式:

$$\boldsymbol{V}'=\boldsymbol{V}\boldsymbol{A}, \tag{6.17}$$

其中 $\boldsymbol{V}'=[T(v_1),\cdots,T(v_n)]$. 同样, (6.13) 式也可表为如下矩阵形式:

$$\boldsymbol{U}'=\boldsymbol{U}\boldsymbol{B}, \tag{6.18}$$

其中 $\boldsymbol{U}'=[T(u_1),\cdots,T(u_n)]$. 若将 T 作用在 (6.14) 中各等式上, 则得

$$T(u_j)=\sum_{k=1}^{n}c_{kj}T(v_k),\quad j=1,2,\cdots,n,$$

将其写成矩阵形式, 则得

$$\boldsymbol{U}'=\boldsymbol{V}'\boldsymbol{C}, \tag{6.19}$$

由 (6.17) 与 (6.19) 两式得 $\boldsymbol{U}'=\boldsymbol{V}\boldsymbol{A}\boldsymbol{C}$, 再由 (6.16) 式即得

$$\boldsymbol{U}'=\boldsymbol{U}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}.$$

将此式与等式 (6.18) 相比较, 即得

$$\boldsymbol{U}\boldsymbol{B}=\boldsymbol{U}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}.$$

由于 $1\times n$ 矩阵 $\boldsymbol{U}\boldsymbol{B}$ 与 $\boldsymbol{U}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 的各个元素都是基元素 $u_1,\cdots,u_n$ 的线性组合, 由 $u_1,\cdots,u_n$ 的线性无关性即得 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 即矩阵等式 (6.12) 成立, 从而我们已经证明了下述定理.

定理 6.8　*如果 $n\times n$ 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 表示同一个线性变换, 则存在非奇异矩阵 $\boldsymbol{C}$ 使得*

$$\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}.$$

而且, 如果 $\boldsymbol{A}$ 是 T 在基 $\boldsymbol{V}=[v_1,\cdots,v_n]$ 下的矩阵, $\boldsymbol{B}$ 是 T 在基 $\boldsymbol{U}=[u_1,\cdots,u_n]$ 下的矩阵, 则可取从基 $\boldsymbol{V}$ 到基 $\boldsymbol{U}$ 的转移矩阵作为 $\boldsymbol{C}$, 即 $\boldsymbol{C}$ 为满足条件 $\boldsymbol{U}=\boldsymbol{V}\boldsymbol{C}$ 的非奇异矩阵. □

定理 6.8 的逆命题也成立.

定理 6.9　*设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为两个 $n\times n$ 矩阵, 若存在 $n\times n$ 非奇异矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可表示同一个线性变换.*

证明 选择 n 维线性空间的一组基 $\boldsymbol{V}=[v_1,\cdots,v_n]$, 设

$$\boldsymbol{C}=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix},$$

令

$$u_j=\sum_{k=1}^{n} c_{kj}v_k, \quad j=1,2,\cdots,n, \tag{6.20}$$

再令 $\boldsymbol{U}=[u_1,\cdots,u_n]$. 由于 $\boldsymbol{C}$ 是非奇异矩阵, 因此 $\boldsymbol{C}$ 表示一个可逆线性变换, 从而 $\boldsymbol{U}=[u_1,\cdots,u_n]$ 也是一组基. $\boldsymbol{V}$ 与 $\boldsymbol{U}$ 这两组基之间的关系也可更简洁地表示为矩阵形式. 首先, 等式 (6.20) 可表示为

$$\boldsymbol{U}=\boldsymbol{V}\boldsymbol{C},$$

由此得

$$\boldsymbol{V}=\boldsymbol{U}\boldsymbol{C}^{-1}.$$

由于 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 我们又有

$$\boldsymbol{U}\boldsymbol{B}=\boldsymbol{U}\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{V}\boldsymbol{A}\boldsymbol{C}, \tag{6.21}$$

上式将帮助我们证明矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 表示同一个线性变换.

令 T 为在基 $\boldsymbol{V}$ 之下以 $\boldsymbol{A}$ 为矩阵的线性变换, S 为在基 $\boldsymbol{U}$ 之下以 $\boldsymbol{B}$ 为矩阵的线性变换, 则

$$T(v_k)=\sum_{i=1}^{n} a_{ik}v_i, \quad k=1,2,\cdots,n. \tag{6.22}$$

$$S(u_j)=\sum_{k=1}^{n} b_{kj}u_k, \quad j=1,2,\cdots,n. \tag{6.23}$$

我们要证明 $S=T$, 而这只要证明对每个基元素 u_j 都有 $S(u_j)=T(u_j)$ 即可. 为此, 我们将 (6.22) 与 (6.23) 两式分别表示为下述简单的矩阵形式:

$$[T(v_1),\cdots,T(v_n)]=\boldsymbol{V}\boldsymbol{A}, \quad [S(u_1),\cdots,S(u_n)]=\boldsymbol{U}\boldsymbol{B}.$$

将线性变换 T 作用在等式 (6.20) 的两边即得 $T(u_j)=\sum\limits_{k=1}^{n} c_{kj}T(v_k)$, 表示成矩阵形式即为

$$[T(u_1),\cdots,T(u_n)]=[T(v_1),\cdots,T(v_n)]\boldsymbol{C}=\boldsymbol{V}\boldsymbol{A}\boldsymbol{C}.$$

而由等式 (6.21) 我们又有 $\boldsymbol{U}\boldsymbol{B}=\boldsymbol{V}\boldsymbol{A}\boldsymbol{C}$, 因此

$$[S(u_1),\cdots,S(u_n)]=[T(u_1),\cdots,T(u_n)],$$

从而对每个 u_j 都有 $S(u_j)=T(u_j)$, $1\leqslant j\leqslant n$. 于是得 $S=T$, 因此矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 表示同一个线性变换. □

定义 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为两个 $n\times n$ 矩阵, 如果存在非奇异矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似** (similar).

结合定理 6.8 与 6.9 即得下述定理.

定理 6.10 两个 $n \times n$ 矩阵相似的充分必要条件是它们表示同一个线性变换. □

相似矩阵有许多共同的性质. 例如, 由于

$$\det(\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}) = \det(\boldsymbol{C}^{-1})(\det \boldsymbol{A})(\det \boldsymbol{C}) = \det(\boldsymbol{A})$$

因此, 相似矩阵有相同的行列式. 由此性质可给出下述定理.

定理 6.11 相似矩阵有相同的特征多项式, 从而有相同的特征值.

证明 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则有非奇异矩阵 $\boldsymbol{C}$, 使得 $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$. 于是得

$$\lambda \boldsymbol{I} - \boldsymbol{B} = \lambda \boldsymbol{I} - \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \lambda \boldsymbol{C}^{-1}\boldsymbol{I}\boldsymbol{C} - \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{C}^{-1}(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{C},$$

即 $\lambda \boldsymbol{I} - \boldsymbol{B}$ 与 $\lambda \boldsymbol{I} - \boldsymbol{A}$ 为相似矩阵, 因此它们的行列式相同, 即

$$\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}).$$ □

结合定理 6.10 与定理 6.11 可知, 表示一个给定线性变换 T 的所有矩阵的特征多项式相同, 因此这个多项式也叫作 T 的**特征多项式**.

结合定理 6.7、定理 6.2 与定理 6.8, 我们可得下述定理. 在这个定理中, F 表示实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$.

定理 6.12 设 T 是将某个 n 维线性空间映入自己的线性变换, F 为纯量域. 若 T 的特征多项式在 F 中有 n 个不同的根 $\lambda_1, \ldots, \lambda_n$, 则我们有:

(a) 分别属于这些特征值的特征向量 $u_1, \cdots, u_n$ 组成此 n 维线性空间的一组基;

(b) T 在有序基 $\boldsymbol{U} = [u_1, \cdots, u_n]$ 下的矩阵是以特征值作为主对角线元素的对角矩阵

$$\Lambda = \operatorname{diag}[\lambda_1, \cdots, \lambda_n];$$

(c) 如果 $\boldsymbol{A}$ 是 T 在另一组基 $\boldsymbol{V} = [v_1, \cdots, v_n]$ 下的矩阵, 则 $\Lambda = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 此处 $\boldsymbol{C}$ 是将基 $\boldsymbol{V}$ 变为基 $\boldsymbol{U}$ 的非奇异转移矩阵, 即

$$\boldsymbol{U} = \boldsymbol{V}\boldsymbol{C}.$$

证明 由定理 6.7, 每个 λ_i 都是 T 的特征值, 因此共有 n 个不同的特征值, 从而由定理 6.2 可知, 它们所对应的特征向量 $u_1, \cdots, u_n$ 线性无关, 因此组成一组基, 即得 (a). 由于 $T(u_i) = \lambda_i u_i$, 因此 T 在基 $\boldsymbol{U}$ 下的矩阵是对角矩阵 Λ, 由此得 (b). 由定理 6.8 即得 (c). □

注 定理 6.12 中的非奇异转移矩阵 $\boldsymbol{C}$ 也叫作**对角化矩阵** (diagonalizing matrix). 如果 $\boldsymbol{V}$ 是由单位坐标向量组成的基 $(\boldsymbol{I}_1, \cdots, \boldsymbol{I}_n)$, 则由定理 6.12 中的等式 $\boldsymbol{U} = \boldsymbol{V}\boldsymbol{C}$ 可知, 矩阵 $\boldsymbol{C}$ 的第 k 列由特征向量 $\boldsymbol{u}_k$ 在基 $(\boldsymbol{I}_1, \cdots, \boldsymbol{I}_n)$ 下的坐标所组成.

如果矩阵 $\boldsymbol{A}$ 的特征值各不相同, 则 $\boldsymbol{A}$ 相似于对角矩阵. 如果 $\boldsymbol{A}$ 有相同的特征值, $\boldsymbol{A}$ 仍然有可能相似于对角矩阵, 不过这一情形当且仅当属于各个特征值的线性无关特征向量的个数等于各个特征值的重数时才会发生. 下述习题中将给出有关的例子.

6.12 习　　题

1. 证明矩阵 $\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$ 与 $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$ 有相同的特征值但不相似.
2. 在下述各小题中, 或者找出一个矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵, 或者证明这样的矩阵 $\boldsymbol{C}$ 不存在.

   (a) $\boldsymbol{A}=\begin{bmatrix}1 & 0\\1 & 3\end{bmatrix}$.　　(b) $\boldsymbol{A}=\begin{bmatrix}1 & 2\\5 & 4\end{bmatrix}$.

   (c) $\boldsymbol{A}=\begin{bmatrix}2 & 1\\-1 & 4\end{bmatrix}$.　　(d) $\boldsymbol{A}=\begin{bmatrix}2 & 1\\-1 & 0\end{bmatrix}$.
3. 给定平面上的三组基, 平面上一点在这三组基下的坐标分别为 (x_1,x_2), (y_1,y_2) 与 (z_1,z_2). 假定 $[y_1,y_2]=[x_1,x_2]\boldsymbol{A}$, $[z_1,z_2]=[x_1,x_2]\boldsymbol{B}$, $[z_1,z_2]=[y_1,y_2]\boldsymbol{C}$, 其中 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 为 2×2 矩阵. 试由 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 表示 $\boldsymbol{C}$.
4. 在下列各小题中, 证明 $\boldsymbol{A}$ 有相同的特征值, 但是 $\boldsymbol{A}$ 仍然有 3 个线性无关的特征向量. 试求非奇异矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵.

   (a) $\boldsymbol{A}=\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix}$.　　(b) $\boldsymbol{A}=\begin{bmatrix}1 & -1 & -1\\1 & 3 & 1\\-1 & -1 & 1\end{bmatrix}$.
5. 证明下面 2 个矩阵都不相似于对角矩阵, 但它们都相似于某个形如 $\begin{bmatrix}\lambda & 0\\1 & \lambda\end{bmatrix}$ 的三角形矩阵, 其中 λ 是 $\boldsymbol{A}$ 的特征值.

   (a) $\begin{bmatrix}2 & -1\\0 & 2\end{bmatrix}$.　　(b) $\begin{bmatrix}2 & 1\\-1 & 4\end{bmatrix}$.
6. 确定矩阵 $\begin{bmatrix}0 & -1 & 0\\0 & 0 & 1\\-1 & -3 & 3\end{bmatrix}$ 的特征值和特征向量并由此证明它不相似于对角矩阵.
7. (a) 证明方阵 $\boldsymbol{A}$ 可逆的充分必要条件为 0 不是 $\boldsymbol{A}$ 的特征值.

   (b) 如果 $\boldsymbol{A}$ 可逆, 证明 $\boldsymbol{A}^{-1}$ 的特征值是 $\boldsymbol{A}$ 的特征值的倒数.
8. 给定 $n\times n$ 实矩阵 $\boldsymbol{A}$ 使 $\boldsymbol{A}^2=-\boldsymbol{I}$. 证明下述各结论:

   (a) $\boldsymbol{A}$ 可逆.　(b) n 为偶数.　(c) $\boldsymbol{A}$ 没有实特征值.　(d) $\det\boldsymbol{A}=1$.
9. 对下列各小题中的命题, 如果你认为是正确的, 则给出证明; 如果你认为是不正确的, 则 (给出一个反例或) 说明为什么它不成立. 除非另有说明, $\boldsymbol{A}$, $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 都是复方阵.

   (a) 如果 $\boldsymbol{A}$ 可以对角化, 则 $\boldsymbol{A}$ 的特征值各不相同.

   (b) 如果 $\boldsymbol{A}$ 的所有特征值都等于 2, 则对于某个非奇异矩阵 $\boldsymbol{B}$ 有 $\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{B}=2\boldsymbol{I}$.

   (c) 如果 $\boldsymbol{A}$ 的所有特征值都是零, 则 $\boldsymbol{A}$ 为零矩阵.

   (d) 如果 $\boldsymbol{A}$ 的所有特征值都是零, 则 $\boldsymbol{A}$ 相似于零矩阵.

   (e) 如果对某个非奇异矩阵 $\boldsymbol{C}$ 有 $\boldsymbol{B}=\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$, 则对同一个矩阵 $\boldsymbol{C}$ 有 $\boldsymbol{B}^3=\boldsymbol{C}^{-1}\boldsymbol{A}^3\boldsymbol{C}$.

   (f) 若 $\boldsymbol{A}$ 为非奇异矩阵, 则它的所有特征值都不等于零.

   (g) 存在没有实特征值的 3×3 实矩阵.

   (h) 矩阵 $\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\0 & 1 & 1\\0 & 0 & 1\end{bmatrix}$ 相似于某个对角矩阵.

6.13 Cayley-Hamilton 定理

Cayley-Hamilton 定理是说每一个方阵都是它的特征多项式的根. W. R. Hamilton 于 1853 年首先对一类特殊的方阵给出了证明. 数年之后, Arthur Cayley (1821—1895) 宣布对所有方阵这个定理都成立, 不过他没有给出证明. 现在关于这个定理已经有许多不同的证明. 我们将用两种等价的形式来叙述这个定理, 一种是用线性算子的形式, 另一种是用矩阵形式. 我们将证明用算子形式叙述的本定理.

定理 6.13 (关于线性变换的 Cayley-Hamilton 定理) *设 V 为 n 维复线性空间, $T:V\to V$ 为给定的线性变换. 设 $\lambda_1,\cdots,\lambda_n$ 为 T 的特征值,*

$$f_T(\lambda)=(\lambda-\lambda_1)\cdots(\lambda-\lambda_n) \tag{6.24}$$

为 T 的特征多项式. 令 $g(T)$ 表示将 $f_T(\lambda)$ 中的 λ 用 T 代替, λ_k 用 $\lambda_k I$ 代替之后所得到的常系数变换, 即

$$g(T)=(T-\lambda_1 I)\cdots(T-\lambda_n I). \tag{6.25}$$

则 $g(T)$ 是零算子, 即 $g(T)$ 将 V 中每一个向量都映为零向量:

$$g(T)(x)=O,\ \text{对所有}\ x\in V.$$

注 每一个特征值 λ_k 都满足多项式方程 $f_T(\lambda)=0$, Cayley-Hamilton 定理则是说 T 满足方程 $f_T(T)=O$.

证明 由于取复数域为纯量域, 因此由定理 6.5, 我们可以取 V 的一组基使 T 在这组基下的矩阵是上三角形矩阵, 将此上三角形矩阵记作 $\boldsymbol{U}$, 则 $\boldsymbol{U}$ 的主对角线元素都是 T 的特征值, 因此 $\boldsymbol{U}$ 可表为下述形式:

$$\boldsymbol{U}=\begin{bmatrix}\lambda_1 & u_{12} & \cdots & u_{1n}\\ 0 & \lambda_2 & \cdots & u_{2n}\\ 0 & 0 & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}.$$

于是算子 $g(T)$ 的矩阵可表为如下乘积形式:

$$(\boldsymbol{U}-\lambda_1\boldsymbol{I})\cdots(\boldsymbol{U}-\lambda_n\boldsymbol{I}). \tag{6.26}$$

我们将通过考察上述乘积中各个因子的作用来证明此乘积将每一个列向量 $\boldsymbol{X}=[x_1,\cdots,x_n]^{\mathrm{T}}$ 变为零向量.

乘积 (6.26) 中最后的因子 $\boldsymbol{U}-\lambda_n\boldsymbol{I}$ 有下述形式:

$$\boldsymbol{U}-\lambda_n\boldsymbol{I}=\begin{bmatrix}\lambda_1-\lambda_n & u_{12} & \cdots & u_{1n}\\ 0 & \lambda_2-\lambda_n & \cdots & u_{2n}\\ 0 & 0 & \ddots & \vdots\\ 0 & 0 & \cdots & 0\end{bmatrix},$$

它的最后一行的各个元素都是零. 这个矩阵与任意一个列向量 $\boldsymbol{X} = [x_1, \cdots, x_n]^{\mathrm{T}}$ 的乘积的最后一个元素必为零, 即

$$(\boldsymbol{U} - \lambda_n \boldsymbol{I})\boldsymbol{X} = (\boldsymbol{U} - \lambda_n \boldsymbol{I}) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} * \\ * \\ \vdots \\ * \\ 0 \end{bmatrix},$$

其中 $*$ 代表该位置上对应的元素. 乘积表达式 (6.26) 中倒数第 2 个因子可表为如下形式:

$$\boldsymbol{U} - \lambda_{n-1}\boldsymbol{I} = \begin{bmatrix} \lambda_1 - \lambda_{n-1} & u_{12} & \cdots & u_{1n} \\ 0 & \lambda_2 - \lambda_{n-1} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & * \\ 0 & \cdots & 0 & \lambda_n - \lambda_{n-1} \end{bmatrix}$$

因此它与任一形如 $(*, *, \cdots, *, 0)^{\mathrm{T}}$ 的列向量的乘积的最后 2 个分量都是零, 即我们有

$$(\boldsymbol{U} - \lambda_{n-1}\boldsymbol{I})(\boldsymbol{U} - \lambda_n \boldsymbol{I})\boldsymbol{X} = \begin{bmatrix} * \\ * \\ \vdots \\ * \\ 0 \\ 0 \end{bmatrix}.$$

如此继续下去, 我们看到, 每一个新因子 $\boldsymbol{U} - \lambda_k \boldsymbol{I}$ 都将使所得到的列向量增加一个零分量, 从而 (6.26) 中所有因子的乘积将列向量 $\boldsymbol{X}$ 变为零向量. 由此我们证明了线性变换形式的 Cayley-Hamilton 定理, 这个证明同时也指出了每个因子对于使 $\boldsymbol{X}$ 变为零向量所起的作用. □

定理 6.13 给出了 Cayley-Hamilton 定理的线性变换形式, 它也可以用矩阵来表述为另一个等价的形式.

推论 6.14 (Cayley-Hamilton 定理的矩阵形式) 设 $\boldsymbol{A}$ 为一个 $n \times n$ 矩阵, 再令

$$f(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 \tag{6.27}$$

为 $\boldsymbol{A}$ 的特征多项式, 则 $f(\boldsymbol{A}) = \boldsymbol{O}$. 换言之, A 满足下述等式:

$$\boldsymbol{A}^n + c_{n-1}\boldsymbol{A}^{n-1} + \cdots + c_1\boldsymbol{A} + c_0\boldsymbol{I} = \boldsymbol{O}. \tag{6.28}$$ □

例 1 2×2 矩阵 $\boldsymbol{A} = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$ 的特征多项式为

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \lambda^2 - 7\lambda + 6.$$

由 Cayley-Hamilton 定理得 $\boldsymbol{A}^2-7\boldsymbol{A}+6\boldsymbol{I}=\boldsymbol{O}$, 由此又有

$$\boldsymbol{A}^2=7\boldsymbol{A}-6\boldsymbol{I}. \tag{6.29}$$

由上式可得到很多推论. 首先, 我们可以用它来计算 $\boldsymbol{A}$ 的所有高次幂. 例如, 我们有

$$\begin{aligned}\boldsymbol{A}^3&=\boldsymbol{A}\boldsymbol{A}^2=\boldsymbol{A}(7\boldsymbol{A}-6\boldsymbol{I})=7\boldsymbol{A}^2-6\boldsymbol{A}=7(7\boldsymbol{A}-6\boldsymbol{I})-6\boldsymbol{A}=43\boldsymbol{A}-42\boldsymbol{I},\\ \boldsymbol{A}^4&=\boldsymbol{A}\boldsymbol{A}^3=\boldsymbol{A}(43\boldsymbol{A}-42\boldsymbol{I})=43\boldsymbol{A}^2-42\boldsymbol{A}=43(7\boldsymbol{A}-6\boldsymbol{I})-42\boldsymbol{A}=259\boldsymbol{A}+258\boldsymbol{I}.\end{aligned}$$

我们也可以利用上述等式求 $\boldsymbol{A}^{-1}$. 将等式 (6.29) 写成 $\boldsymbol{A}(\boldsymbol{A}-7\boldsymbol{I})=-6\boldsymbol{I}$ 的形式, 再在两边同乘以 $-\frac{1}{6}$, 则得 $\boldsymbol{A}\{\frac{1}{6}(7\boldsymbol{I}-\boldsymbol{A})\}=\boldsymbol{I}$. 因此 $\boldsymbol{A}^{-1}=\frac{1}{6}(7\boldsymbol{I}-\boldsymbol{A})$. □

例 2　3×3 矩阵 $\boldsymbol{A}=\begin{bmatrix}5&4&0\\1&2&0\\1&2&2\end{bmatrix}$ 的特征多项式为

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^3-9\lambda^2+20\lambda-12=(\lambda-1)(\lambda-2)(\lambda-6),$$

因此 $\boldsymbol{A}$ 满足等式

$$\boldsymbol{A}^3-9\boldsymbol{A}^2+20\boldsymbol{A}-12\boldsymbol{I}=\boldsymbol{O}. \tag{6.30}$$

上述等式可用来将 $\boldsymbol{A}^3$ 以及 $\boldsymbol{A}$ 的更高次幂表示成 $\boldsymbol{I}$, $\boldsymbol{A}$ 与 $\boldsymbol{A}^2$ 的线性组合, 例如我们有

$$\begin{aligned}\boldsymbol{A}^3&=9\boldsymbol{A}^2-20\boldsymbol{A}+12\boldsymbol{I},\\ \boldsymbol{A}^4&=9\boldsymbol{A}^3-20\boldsymbol{A}^2+12\boldsymbol{A}=9(9\boldsymbol{A}^2-20\boldsymbol{A}+12\boldsymbol{I})-20\boldsymbol{A}^2+12\boldsymbol{A}\\ &=61\boldsymbol{A}^2-168\boldsymbol{A}+108\boldsymbol{I}.\end{aligned}$$

它也可以用来将 $\boldsymbol{A}^{-1}$ 表示为 $\boldsymbol{A}$ 的多项式. 由等式 (6.30) 我们有

$$\boldsymbol{A}(\boldsymbol{A}^2-9\boldsymbol{A}+20\boldsymbol{I})=12\boldsymbol{I},$$

从而 $\boldsymbol{A}^{-1}=\frac{1}{12}(\boldsymbol{A}^2-9\boldsymbol{A}+20\boldsymbol{I})$. □

6.14　习　　题

在习题 1~4 中, 将 $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^2$ 及 $\boldsymbol{A}$ 的更高次幂表示为 $\boldsymbol{I}$ 与 $\boldsymbol{A}$ 的线性组合.

1. $\boldsymbol{A}=\begin{bmatrix}1&0\\1&1\end{bmatrix}$.　　2. $\boldsymbol{A}=\begin{bmatrix}1&0\\1&2\end{bmatrix}$.

3. $\boldsymbol{A}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$.　　4. $\boldsymbol{A}=\begin{bmatrix}-1&0\\0&1\end{bmatrix}$.

在习题 5~7 中, 计算 $\boldsymbol{A}^n$, 并将 $\boldsymbol{A}^3$ 用 $\boldsymbol{I}$, $\boldsymbol{A}$ 和 $\boldsymbol{A}^2$ 表示.

5. $\boldsymbol{A}=\begin{bmatrix}0&1&1\\0&0&1\\0&0&0\end{bmatrix}$.　　6. $\boldsymbol{A}=\begin{bmatrix}0&1&1\\0&1&1\\0&0&0\end{bmatrix}$.　　7. $\boldsymbol{A}=\begin{bmatrix}2&0&0\\0&1&0\\0&1&1\end{bmatrix}$.

6.15 Jordan 标准型

由定理 6.5, 复数域上有限维线性空间到它自身内的任一线性变换都可以三角化. 本节将证明存在这样的一种三角形矩阵表示, 它能够分解成称为 **Jordan 块** (block) 的特殊三角形矩阵. 下面给出 2 阶、3 阶与 4 阶 Jordan 块的范例:

$$\boldsymbol{J}=\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix},\quad \boldsymbol{J}=\begin{bmatrix}\mu & 1 & 0\\ 0 & \mu & 1\\ 0 & 0 & \mu\end{bmatrix},\quad \boldsymbol{J}=\begin{bmatrix}\nu & 1 & 0 & 0\\ 0 & \nu & 1 & 0\\ 0 & 0 & \nu & 1\\ 0 & 0 & 0 & \nu\end{bmatrix}$$

定义 具有下述性质的上三角形矩阵 $[a_{ij}]$ 都叫做 **Jordan 块**: 所有的主对角线元素 a_{ii} 都相等, 所有主对角线上方的元素 $a_{i,i+1}$ 都等于 1, 其余元素都等于零. 所有的 1×1 矩阵也叫作 Jordan 块.

下述定理刻划出了 Jordan 块的全部特征值和特征向量.

定理 6.15 设 $\boldsymbol{J}$ 是一个 k 阶 Jordan 块, 则 $\boldsymbol{J}$ 只有一个特征值, 它等于 J 的主对角线元素, 与其对应的特征向量是 k 维单位坐标向量 $[1,0,\cdots,0]$ 的非零纯量倍.

证明 设 $\boldsymbol{J}$ 的主对角线元素为 λ, 则列向量 $\boldsymbol{X}=[x_1,x_2,\cdots,x_k]^{\mathrm{T}}$ 满足矩阵方程 $\boldsymbol{JX}=\lambda\boldsymbol{X}$ 的充分必要条件是 $\boldsymbol{X}$ 的各分量满足下述 k 个纯量方程:

$$\begin{aligned}\lambda x_1+x_2&=\lambda x_1,\\ \lambda x_2+x_3&=\lambda x_2,\\ &\ \vdots\\ \lambda x_{k-1}+x_k&=\lambda x_{k-1},\\ \lambda x_k&=\lambda x_k.\end{aligned}$$

由前 $k-1$ 个方程得

$$x_2=x_3=\cdots=x_k=0,$$

因此 λ 是 $\boldsymbol{J}$ 的一个特征值, 而所有特征向量都可表示为 $x_1[1,0,\cdots,0]$ 的形式, 其中 $x_1\neq 0$.

为证明 λ 是 $\boldsymbol{J}$ 所仅有的特征值, 假定对某个 $\mu\neq\lambda$ 有 $\boldsymbol{JX}=\mu\boldsymbol{X}$, 则 $\boldsymbol{X}$ 的各分量满足下述 k 个纯量方程:

$$\begin{aligned}\lambda x_1+x_2&=\mu x_1,\\ \lambda x_2+x_3&=\mu x_2,\\ &\ \vdots\\ \lambda x_{k-1}+x_k&=\mu x_{k-1},\\ \lambda x_k&=\mu x_k.\end{aligned}$$

由于 $\mu \neq \lambda$, 因此由最后的方程得 $x_k = 0$, 从而再依次由第 $k-1$, 第 $k-2$ 直到第 1 个方程得 $x_{k-1} = \cdots = x_2 = x_1 = 0$, 因此只有零向量才能满足 $\boldsymbol{JX} = \mu \boldsymbol{X}$, 从而 μ 不可能是 J 的特征值, 亦即 λ 是 J 所仅有的特征值. □

下述定理显示了 Jordan 块的重要性.

定理 6.16 (Jordan 标准型)　*设 V 为 n 维复线性空间, $T: V \to V$ 是 V 到其自身内的一个线性变换, 则存在 V 的一组基, 使 T 在这组基下的矩阵是一个分块对角矩阵 $\operatorname{diag}(J_1, \cdots, J_m)$, 其中各个 J_k 都是 Jordan 块.*

证明　对维数 n 作归纳法. 当 $n=1$ 时结论显然成立. 假设结论对作用在维数 $< n$ 的线性空间上的所有线性变换都成立, 我们要证明结论对 n 维空间也成立.

令 λ 为 T 的一个取定的特征值 (由定理 6.4 知 T 的特征值存在), 再令 $S = T - \lambda \boldsymbol{I}$, 若能证明 S 有一个 Jordan 型的矩阵表示, 则 $T = S + \lambda \boldsymbol{I}$ 也有这样的表示.

由于 S 将 T 的一个特征向量映为 O, 因此 S 的**零化度** $\nu \geqslant 1$, 秩 $r = n - \nu < n$, 此处 $\nu = \dim N(S)$ 为 S 的零空间的维数, r 为值域 $S(V)$ 的维数.

证明方案是: 令 S' 为线性变换 S 在子空间 $S(V)$ 上的限制, 由于 $r = \dim S(V) < n$, 因此由归纳假设, 存在 $S(V)$ 的一组基 E, 使 S' 在这组基下的矩阵是 Jordan 型矩阵. 然后我们证明 E 可扩充成 V 的一组基 E', 使得 S 在 E' 下的矩阵是 Jordan 型矩阵.

由于 S' 将 $S(V)$ 映入 $S(V)$ 本身, 因此由归纳假设, 存在 $S(V)$ 的一组由 r 个元素组成的基 E, 使得 S' 在这组基下的矩阵为 Jordan 型矩阵, 设为

$$m_E(S') = \operatorname{diag}(\boldsymbol{J}_1, \cdots, \boldsymbol{J}_m),$$

其中 $\boldsymbol{J}_i$ 是一个 n_i 阶 Jordan 块, 并且各 Jordan 块的阶数之和为 $n_1 + \cdots + n_m = r$. 将 E 中的向量记作 $e_{i,j}$, $1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n_i$, 并将它们分成 m 个互不相交的有序子集 $E_1, E_2, \cdots, E_m$, 它们分别对应于 m 个 Jordan 块, 不妨设

$$E_i = \{e_{i,1}, e_{i,2}, \cdots, e_{i,n_i}\}, \quad i = 1, 2, \cdots, m.$$

若 Jordan 块 $\boldsymbol{J}_i$ 的阶数 $n_i = 1$ 且其唯一的元素为 λ_i, 则我们有

$$S(e_{i,1}) = \lambda_i e_{i,1}. \tag{6.31}$$

若 J_i 的阶 $n_i > 1$, 且其主对角线元素为 λ_i, 则等式 (6.31) 对 $e_{i,1}$ 成立, 而对 E_i 的其余元素我们有

$$S(e_{i,j}) = e_{i,j-1} + \lambda_i e_{i,j}, \quad j = 2, \cdots, n_i. \tag{6.32}$$

某些特征值 λ_i 可能为零. 如果出现这种情形, 则在必要时适当对基 E 中的元素重新排列可使这些零特征值对应的 Jordan 块为 $\boldsymbol{J}_1, \cdots, \boldsymbol{J}_q$. 于是当 $i \leqslant q$ 时 $\boldsymbol{J}_i$ 的秩为 $n_i - 1$, 当 $i > q$ 时 J_i 的秩为 n_i. 由于 $m_E(S')$ 的秩为各 $\boldsymbol{J}_i$ 的秩的和, 因此

$$\operatorname{rank} S' = \operatorname{rank} m_E(S') = \sum_{i=1}^{q}(n_i - 1) + \sum_{i=q+1}^{m} n_i = \sum_{i=1}^{m} n_i - q = r - q.$$

因此 $q = r - \mathrm{rank}S' = S'$ 的零化度. 然而由等式 (6.31) 可知对 $i = 1, \cdots, q$ 都有 $S'(e_{i,1}) = 0$, 因此 $e_{1,1}, \cdots, e_{q,1}$ 这 q 个向量组成零空间 $N(S')$ 的一组基.

由 $N(S') \subseteq N(S)$ 得 $q = \dim N(S') \leqslant \dim N(S) = \nu$. 令 $k = \nu - q$, 在 $N(S)$ 中存在 k 个元素 $w_1, \cdots, w_k$ 使得下述元素

$$\{e_{1,1}, e_{2,1}, \cdots, e_{q,1}, w_1, \cdots, w_k\}$$

组成 $N(S)$ 的一组基. 对 $i = 1, 2, \cdots, q$, 令 v_i 表示 V 中使得

$$S(v_i) = e_{i,n_i}, \tag{6.33}$$

的元素. 我们要证明元素组

$$E' = E \cup \{w_1, \cdots, w_k\} \cup \{v_1, \cdots, v_q\}$$

组成 V 的一组基.

我们先证明 E' 线性无关, 为此假定 E' 中元素的某个线性组合表示零元素, 不妨设

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} c_{ij}e_{i,j} + \sum_{i=1}^{k} a_i w_i + \sum_{i=1}^{q} b_i v_i = O. \tag{6.34}$$

将线性变换 S 作用于上式, 再由 $S(w_i) = O$ 及 (6.31), (6.32) 与 (6.33) 各式可得

$$\sum_{i=1}^{m} c_{i1}\lambda_i e_{i,1} + \sum_{i=1}^{m}\sum_{j=2}^{n_i} c_{ij}\{e_{i,j-1} + \lambda_i e_{i,j}\} + \sum_{i=1}^{q} b_i e_{i,n_i} = O. \tag{6.35}$$

由于当 $i \leqslant q$ 时, $\lambda_i = 0$, 因此式 (6.35) 可改写如下:

$$\sum_{i=q+1}^{m} c_{i1}\lambda_i e_{i,1} + \sum_{i=1}^{m}\sum_{j=2}^{n_i} c_{ij}e_{i,j-1} + \sum_{i=q+1}^{m}\sum_{j=2}^{n_i} c_{ij}\lambda_i e_{i,j} + \sum_{i=1}^{q} b_i e_{i,n_i} = O. \tag{6.36}$$

由于上式最后一个和式中的 $i \leqslant q$, 因此各个元素 e_{i,n_j} 不在其他和式中出现, 因此由线性无关性可知 $b_i = 0, i = 1, 2, \cdots, q$. 将此代入式 (6.34) 得

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} c_{ij}e_{i,j} = -\sum_{i=1}^{k} a_i w_i. \tag{6.37}$$

上式左边是 $S(V)$ 中元素, 右边是 $N(S)$ 中元素, 因此它们都属于 $S(V) \cap N(S) = N(S')$, 从而都由 $\{e_{1,1}, e_{2,1}, \cdots, e_{q,1}\}$ 生成. 于是对 $j \geqslant 2$ 都有 $c_{ij} = 0$. 再由式 (6.35) 可得, 对每个 $i \geqslant q+1$ 都有 $c_{i1} = 0$. 从而等式 (6.37) 变为

$$\sum_{i=1}^{q} c_{i1}e_{i,1} + \sum_{i=1}^{k} a_i w_i = O.$$

由于 $\{e_{1,1}, e_{2,1}, \cdots, e_{q,1}, w_1, \cdots, w_k\}$ 是 $N(S)$ 的一组基, 因此对所有 i 都有 $c_{i1} = a_i = 0$, 由此证得 E' 的线性无关性. E' 中的元素的个数是 $r + k + q = r + \nu = n$, 因此 E' 是 V 的一组基.

线性变换 S 依照 (6.31) 与 (6.32) 两式作用在基 E 的元素上. 对 $1 \leqslant i \leqslant k$, S 将 w_i 映为零元素, 对 $i \leqslant q$, S 将 v_i 映为 e_{i,n_i}. 于是对 $i = 1, \cdots, q$, 如果我们将 v_i 记作 e_{i,n_i+1}, 则由于当 $i \leqslant q$ 时有 $\lambda_i = 0$, 因此等式 (6.32) 对 $j = n_i + 1$ 仍然成立. 对 $i = 1, 2, \cdots, q$, 我们令

$$E_i' = \{e_{i,1}, \cdots, e_{i,n_i}, e_{i,n_i+1}\},$$

如果我们依下述方案重新排列 E' 中元素的次序:

$$E' = \{w_1, \cdots, w_k\} \cup E_1' \cup \cdots \cup E_q' \cup E_{q+1} \cup \cdots \cup E_m,$$

则 S 在基 E' 下的矩阵为 Jordan 标准型, 即得定理的证明. □

6.16　关于特征值与特征向量的综合性习题

1. (a) 求矩阵 $\boldsymbol{A} = \begin{bmatrix} -1 & -10 & -19 \\ 0 & 4 & 8 \\ 0 & 0 & 0 \end{bmatrix}$ 的特征值与特征向量.

   (b) 求一个元素为整数的可逆矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{-1}\boldsymbol{AC}$ 为对角矩阵.

2. 验证矩阵 $\boldsymbol{A} = \begin{bmatrix} 5 & 4 & 3 \\ -1 & 0 & -3 \\ 1 & -2 & 1 \end{bmatrix}$ 有特征值 $-2, 4, 4$. 由定理 6.16, 存在可逆矩阵 $\boldsymbol{C}$ 将 $\boldsymbol{A}$ 化为 Jordan 标准型:

$$\boldsymbol{C}^{-1}\boldsymbol{AC} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{bmatrix}.$$

   (矩阵 $\boldsymbol{C}$ 不是唯一的, 因若将 $\boldsymbol{C}$ 用它的任一非零纯量倍数代替, 上式仍然成立). 试确定这样的一个可逆矩阵 $\boldsymbol{C}$ 使 $c_{11} = 1$. [**建议**: $\boldsymbol{A}$ 的特征向量可用来确定 $\boldsymbol{C}$ 的两列. 再用待定系数法确定 $\boldsymbol{C}$ 的第 3 列向量, 使得 $\boldsymbol{AC} = \boldsymbol{CB}$, 此处 $\boldsymbol{B}$ 为 Jordan 标准型].

3. (a) 验证矩阵 $\boldsymbol{A} = \begin{bmatrix} -11 & -7 & -5 \\ 16 & 11 & 6 \\ 12 & 6 & 7 \end{bmatrix}$ 有特征值 1, 3, 3.

   (b) 求可逆矩阵 $\boldsymbol{C}$ 使 $c_{11} = 1$ 并且 $\boldsymbol{C}$ 将 $\boldsymbol{A}$ 化为 Jordan 标准型:

$$\boldsymbol{C}^{-1}\boldsymbol{AC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

4. (a) 验证矩阵 $\boldsymbol{A} = \begin{bmatrix} -1 & 3 & 0 \\ 0 & 2 & 0 \\ 2 & 1 & -1 \end{bmatrix}$ 有特征值 $2, -1, -1$.

   (b) 求可逆矩阵 $\boldsymbol{C}$ 使 $c_{11} = 1$ 并且 $\boldsymbol{C}$ 将 $\boldsymbol{A}$ 化为如下 Jordan 标准型:

$$\boldsymbol{C}^{-1}\boldsymbol{AC} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$

   (c) 求所有可逆矩阵 $\boldsymbol{D}$ 使 $\boldsymbol{D}$ 将 $\boldsymbol{A}$ 化为如下 Jordan 标准型:

$$D^{-1}AD = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

5. 验证矩阵 $\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ 有两个复特征值并找出一对正交的特征向量.
6. 试求一个 3×3 矩阵, 使它的特征值为 1, 2, 3, 并且对应的特征向量依次为 $(1,2,3)^{\mathrm{T}}$, $(1,2,1)^{\mathrm{T}}$ 与 $(3,2,1)^{\mathrm{T}}$.
7. 令 $\boldsymbol{J}$ 表示所有元素都等于 1 的 $n\times n$ 矩阵.
   (a) 用 $\boldsymbol{J}$ 表示 $\boldsymbol{J}^2, \boldsymbol{J}^3, \cdots, \boldsymbol{J}^n$.
   (b) 证明 $\boldsymbol{J}$ 有一个特征值为 n, 其余的特征值都是零.
   (c) 当 $n=4$ 时, 求 $\boldsymbol{J}$ 的一组互相正交的特征向量, 其中的一个是 $(1,1,1,1)^{\mathrm{T}}$.
   (d) 对任意纯量 t, 求矩阵 $t\boldsymbol{J}$ 的特征值.
8. 给定一个主对角线元素为 a 而其余元素都等于 t 的 $n\times n$ 矩阵, 证明除了一个例外, 它的其余特征值都等于 $a-t$, 并将这个剩下的特征值用 a 与 t 表示.
9. 给定一个 $n\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$, 其中, 当 $i+j=n+1$ 时, $a_{ij}=n$, 在其他情形都有 $a_{ij}=0$, 例如在 $n=2$ 时我们有 $\boldsymbol{A}=\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$, 在 $n=3$ 时, 有 $\boldsymbol{A}=\begin{bmatrix} 0 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{bmatrix}$. 试用 n, $\boldsymbol{A}$ 及单位矩阵 $\boldsymbol{I}$ 来表示 $\boldsymbol{A}^2$, $\boldsymbol{A}^3$ 及 $\boldsymbol{A}^{-1}$.
10. 令 $\boldsymbol{A}=\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.
    (a) 对每个整数 $n\geqslant 2$, 找出 (依赖于 n) 纯量 a 与 b 使 $\boldsymbol{A}^n = a\boldsymbol{I}+b\boldsymbol{A}$.
    (b) 确定 $\boldsymbol{A}$ 的所有特征值 (包括重数).
    (c) 求可逆矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵或者说明为什么这样的矩阵 $\boldsymbol{C}$ 不存在.
11. 令 $\boldsymbol{A}=\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$.
    (a) 求 $\boldsymbol{A}^4$.
    (b) 求 $\boldsymbol{A}$ 的特征多项式及所有特征值 (包括重数).
12. 令 $\boldsymbol{A}=(a_{ij})$ 为主对角线元素都是 0 而其余元素都是 1 的 4×4 矩阵, 即当 $i=j$ 时 $a_{ij}=0$, 当 $i\neq j$ 时 $a_{ij}=1$.
    (a) 求常数 a 与 b 使得 $\boldsymbol{A}^2=a\boldsymbol{A}+b\boldsymbol{I}$.
    (b) 证明 $\boldsymbol{A}$ 可逆并求 $\boldsymbol{A}^{-1}$.
    (c) 求出 $\boldsymbol{A}$ 的全体特征值 (包括重数).
    (d) 求出 $\boldsymbol{A}$ 的 4 个线性无关的特征向量或证明它们不存在.

第 7 章　欧氏空间中线性变换的特征值

7.1　特征值与内积

本章研究作用在欧氏空间上的线性变换的特征值与特征向量的性质. 欧氏空间就是定义了一个内积的线性空间, 我们先来回顾内积的基本性质.

在一个**实欧氏空间**中, 两个元素 x 与 y 的**内积**是具有下列性质的一个实数 (x,y):

(1) $(x,y)=(y,x)$　　(对称性),

(2) $(x,\,y+z)=(x,y)+(x,z)$　　(线性性),

(3) $c(x,y)=(cx,y)$　　(齐次性),

(4) $(x,x)>0$, 若 $x\neq 0$　　(正定性).

在**复线性空间**中, 两个元素 x 与 y 的内积 (x,y) 是一个复数, 它同样具有上述各条性质, 只不过把对称性换为下面的**Hermite 对称性**:

(1′)　　$(x,y)=\overline{(y,x)}$,

此处 $\overline{(y,x)}$ 表示 (y,x) 的**共轭复数**. 在性质 (3) 中, 纯量 c 是实数或复数. 由性质 (3) 与 (1′) 我们得到如下关系式:

(3′)　　$(x,cy)=\overline{c}(x,y)$,

它是说当将 y 的系数 c 提到内积符号前面时要取它的共轭复数. 在性质 (1′) 中令 $x=y$ 则可知 (x,x) 是实数, 因此对复线性空间中定义的内积, 性质 (4) 仍然成立.

今后如不作特别说明, 当我们用到“**欧氏空间**”这一条术语时, 这个空间可以是实的也可以是复的. 尽管在我们的应用中, 大多数空间都是有限维的, 但我们在开始时并不作这个限制.

本章的第一个定理指出, 若特征值存在, 则可以用内积表示.

定理 7.1　设 E 为一个欧氏空间, V 为 E 的子空间, $T:V\longrightarrow E$ 为一个线性变换, λ 是 T 的一个特征值, x 是属于 λ 的特征向量, 则

$$\lambda=\frac{(T(x),x)}{(x,x)}.\tag{7.1}$$

证明　由于 $T(x)=\lambda x$, 因此我们有

$$(T(x),x)=(\lambda x,x)=\lambda(x,x),$$

而因 $x\neq O$, 再将两边同除以 (x,x) 即得等式 (7.1).　□

从等式 (7.1) 容易推出关于内积的若干性质, 例如, 由内积的 Hermite 对称性, 我们得到关于 λ 的共轭复数的相伴公式:

$$\overline{\lambda}=\frac{(x,T(x))}{(x,x)}, \tag{7.2}$$

由 (7.1) 和 (7.2) 两式可知 λ 为实数 (即 $\lambda=\overline{\lambda}$) 的充分必要条件是 $(T(x),x)$ 为实数, 即等式

$$(T(x),x)=(x,T(x))$$

对所有属于 λ 的特征向量 x 都成立. (这个条件对实欧氏空间自然满足.)

我们也知道, 特征值 λ 为纯虚数 (即 $\lambda=-\overline{\lambda}$) 的充分必要条件是 $(T(x),x)$ 为纯虚数, 即等式

$$(T(x),x)=-(x,T(x))$$

对所有属于 λ 的特征向量 x 都成立.

7.2 Hermite 变换与斜 Hermite 变换

在本节中我们介绍作用在欧氏空间上的两类重要的线性变换. 这两类线性变换依它们所作用的是实内积空间还是复内积空间而有不同的名字. 在实内积空间情形的变换叫作**对称变换**或**斜对称变换**; 在复内积空间情形的变换叫作 **Hermite 变换**或**斜 Hermite 变换**. 这两类变换在很多不同的应用场合出现. 例如, 在无线维空间中的 Hermite 变换在量子力学中起着重要作用. 我们将主要讨论复空间的情形, 因为这样做并不增加难度.

在本节与下节中, E 表示一个欧氏空间, V 是 E 的子空间, 并假定 E 与 V 都是有限维空间.

定义 设 $T:V\longrightarrow E$ 为线性变换, 如果对 V 中所有的 x 与 y 都有

$$(T(x),y)=(x,T(y)),$$

则称 T 为 V 上的一个 Hermite 变换; 如果对 V 中所有的 x 与 y 都有

$$(T(x),y)=-(x,T(y)),$$

则称 T 为 V 上的一个斜 Hermite 变换.

换言之, 一个 Hermite 变换 T 可以从内积中的一个元素移到另一个元素而不改变这个内积的值, 而当将一个斜 Hermite 变换 T 从一个元素移到另一个元素时这个内积要改变正负号.

注 我们已经说过, 若 E 是实欧氏空间, 则 Hermite 变换叫作对称变换, 斜 Hermite 变换叫作斜对称变换.

关于 Hermite 变换与斜 Hermite 变换的特征值, 我们有下述定理.

定理 7.2 假定 T 有特征值 λ.

(a) 若 T 是 Hermite 变换, 则 λ 是实数: $\lambda = \overline{\lambda}$.

(b) 若 T 是斜 Hermite 变换, 则 λ 是纯虚数: $\lambda = -\overline{\lambda}$.

证明 设 x 为属于 λ 的一个特征向量, 则我们有

$$\lambda = \frac{(T(x), x)}{(x, x)} \quad 且 \quad \overline{\lambda} = \frac{(x, T(x))}{(x, x)}.$$

若 T 为 Hermite 变换, 则有 $(T(x), x) = (x, T(x))$, 从而得 $\lambda = \overline{\lambda}$. 而若 T 为斜 Hermite 变换, 则有 $(T(x), x) = -(x, T(x))$, 从而得 $\lambda = -\overline{\lambda}$. □

注 当 T 是对称变换时, 由定理 7.1 可知若内积是实数则所有的特征值必定都是实数, 因此定理 7.2 并没有告诉我们新内容; 但若 T 是斜对称的, 则 T 的特征值必须既是实数又是纯虚数, 因此斜对称变换的所有特征值 (如果存在) 必定为零.

7.3 属于不同特征值的特征向量的正交性

我们由定理 6.2 知道任一线性变换的属于不同特征值的特征向量是线性无关的. 对于 Hermite 变换和斜 Hermite 变换来说还有进一步的结果.

定理 7.3 设 T 为 Hermite 变换或斜 Hermite 变换, 令 x 与 y 分别为属于 T 的两个不同的特征值 λ 与 μ 的特征向量, 则 x 与 y 正交, 即 $(x, y) = 0$.

证明 由定理的条件知 $T(x) = \lambda x$, $T(y) = \mu y$, 计算两个内积 $(T(x), y)$ 与 $(x, T(y))$, 我们有

$$(T(x), y) = (\lambda x, y) = \lambda(x, y) \quad 及 \quad (x, T(y)) = (x, \mu y) = \overline{\mu}(x, y).$$

若 T 是 Hermite 变换, 则由上式及 $\mu = \overline{\mu}$ 得 $\lambda(x, y) = \overline{\mu}(x, y) = \mu(x, y)$, 从而由 $\lambda \neq \mu$ 得 $(x, y) = 0$; 若 T 是斜 Hermite 变换, 则得 $\lambda(x, y) = -\overline{\mu}(x, y) = \mu(x, y)$, 由此也得 $(x, y) = 0$. □

关于作用在函数空间上的对称变换和斜对称变换的例子将在本章末尾的 7.21 节中给出. (它们需要一点微积分的知识.)

7.4 习题

1. 设 E 为一个欧氏空间, V 为它的子空间, 令 $T: V \longrightarrow E$ 为给定的一个线性变换. 证明: 纯量 λ 是 T 的一个以 x 为特征向量的特征值的充分必要条件是 $x \neq O$ 且对 E 中每一个 y 都有 $(T(x), y) = \lambda(x, y)$.
2. 对线性空间 V 中每一个元素 x, 令 $T(x) = cx$, 其中 c 是一个取定的纯量. 证明: 若 V 是实欧氏空间, 则 T 是对称变换.
3. 设 $T: V \longrightarrow V$ 为一个 Hermite 变换.
   (a) 证明: 对每一个整数 $n > 0$, T^n 也是一个 Hermite 变换, 并且当 T 可逆时 T^{-1} 也是 Hermite 变换.
   (b) 当 T 是斜 Hermite 变换时你能得出关于 T^n 和 T^{-1} 的什么结论?

4. 设 $T_1: V \longrightarrow E$ 与 $T_2: V \longrightarrow E$ 为两个 Hermite 变换.
   (a) 证明：对所有实数 a 与 b, $aT_1 + bT_2$ 都是 Hermite 变换.
   (b) 证明：若 T_1 与 T_2 可交换, 即若 $T_1T_2 = T_2T_1$, 则其乘积 T_1T_2 也是 Hermite 变换.
5. 令 $V = \mathbb{R}^3$ 并以通常的点积作为内积. 令 T 为关于 xy 平面的一个反射变换, 即 $T(i) = i$, $T(j) = j$, $T(k) = -k$. 证明：T 是一个对称变换.
6. 设 V 为某个复欧氏空间 E 的子空间, 令 $T: V \longrightarrow E$ 为一个线性变换, 定义 V 上的一个数值函数 Q 如下：

$$Q(x) = (T(x), x), \quad \text{对} V \text{中所有} x.$$

   (a) 证明：若 T 是 V 上的一个 Hermite 变换, 则对所有 x, $Q(x)$ 都是实数.
   (b) 证明：若 T 是斜 Hermite 变换, 则对所有 x, $Q(x)$ 都是纯虚数.
   (c) 证明：对任意复数 t 都有 $Q(tx) = |t|^2 Q(x)$.
   (d) 证明 $Q(x+y) = Q(x) + Q(y) + (T(x), y) + (T(y), x)$, 并对 $Q(x+ty)$ 给出一个相应的公式.
   (e) 假定 $E = V$, 若对所有 x 都有 $Q(x) = 0$, 证明：对所有 x 都有 $T(x) = O$.
   (f) 证明：若对所有 x, $Q(x)$ 都是实数, 则 T 是 Hermite 变换. [提示：对任意复数 t, $Q(x+ty)$ 都等于它的共轭复数.]

7.5 有限维空间中 Hermite 变换和斜 Hermite 变换的标准正交特征向量组的存在性

定理 7.2 和定理 7.3 都是建立在 T 有一个特征值的假设的基础上的. 如我们所知, 有些变换 T 未必有特征值, 不过, 如果 T 作用在一个有限维复空间上, 则 (由定理 6.4) 特征值总存在, 而且, 特征值就是特征多项式的根. 若 T 是 Hermite 变换, 则所有特征值都是实数; 若 T 是斜 Hermite 变换, 则所有特征值都是纯虚数.

我们在 7.3 节中刚证明过, 当 T 为 Hermite 变换或斜 Hermite 变换时, 属于不同特征值的特征向量互相正交. 利用这一性质可以证明, T 有一个标准正交特征向量组, 它生成整个空间. 若一个正交向量组中每一个向量的范数都是 1, 则称它为**标准正交组**.

定理 7.4 设 V 为 n 维复欧氏空间, $T: V \longrightarrow V$ 为 Hermite 变换或斜 Hermite 变换, 则存在 T 的 n 个特征向量 $u_1, \cdots, u_n$, 它们组成 V 的一组标准正交基, 从而 T 关于这组基的矩阵是对角矩阵 $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, 其中 λ_k 是对应于特征向量 u_k 的特征值.

证明 对维数 n 作归纳法. 若 $n = 1$, 则 T 只有一个特征值, 并且任一范数为 1 的特征向量 u_1 都组成 V 的一组标准正交基.

今设定理对每一个 $n-1$ 维欧氏空间都成立. 为证明定理对 V 也成立, 我们选取 T 的一个特征值 λ_1 以及属于 λ_1 的范数为 1 的一个特征向量 u_1, 于是 $T(u_1) = \lambda_1 u_1$ 并且 $\|u_1\| = 1$. 令 S 为由 u_1 生成的子空间, 我们令 $S^{\perp}$ 为由 V 中所有与 u_1 正交的元素组成的子空间, 即

$$S^{\perp} = \{x : x \in V \text{ 且 } (x, u_1) = 0\}.$$

我们要对 $S^{\perp}$ 应用归纳假设, 为此需要证明 $\dim S^{\perp} = n-1$ 并且 T 将 $S^{\perp}$ 映入 $S^{\perp}$.

由定理 3.7(a) 可知, u_1 是 V 的某组基中的一个元素, 设这组基为 $(u_1, v_2, \cdots, v_n)$, 不失一般性, 我们可假定这是一组标准正交基. (若不然, 可以用 Gram-Schmidt 方法将它变为一组标准正交基, 并且仍以 u_1 作为第一个基元素). 从 $S^{\perp}$ 中任取一个元素 x 并将它表成这组基元素的线性组合, 设

$$x = x_1 u_1 + x_2 v_2 + \cdots + x_n v_n.$$

取 x 与 u_1 的内积, 由于 $(u_1, v_2, \cdots, v_n)$ 是标准正交基, 因此 $x_1 = (x, u_1) = 0$, 从而 x 是由 $v_2, \cdots, v_n$ 生成的子空间中的元素, 即得 $\dim S^{\perp} = n-1$.

下面我们证明 T 将 $S^{\perp}$ 映入它自身. 假设 T 是 Hermite 变换, 若 $x \in S^{\perp}$, 则有

$$(T(x), u_1) = (x, T(u_1)) = (x, \lambda_1 u_1) = \lambda_1 (x, u_1) = 0,$$

因此 $T(x) \in S^{\perp}$, 从而 T 是 $S^{\perp}$ 上的 Hermite 变换. 于是由归纳假设, 存在 T 的 $n-1$ 个特征向量 $u_2, \cdots, u_n$, 它们组成 $S^{\perp}$ 的标准正交基, 从而正交的特征向量组 $u_1, \cdots, u_n$ 是 V 的一组标准正交基, 因此当 T 为 Hermite 变换时定理得证. 同理可证当 T 为斜 Hermite 变换时定理也成立. □

7.6 Hermite 变换与斜 Hermite 变换的矩阵表示

在本节中我们仍假定 V 是有限维欧氏空间. 一个 Hermite 变换或斜 Hermite 变换可以用它在任何一组基元素上的作用来刻划.

定理 7.5 设 $(u_1, \cdots, u_n)$ 为 V 的一组基, $T : V \longrightarrow V$ 为一个线性变换. 我们有:

(a) 当且仅当对所有 i 与 j 都有 $(T(u_j), u_i) = (u_j, T(u_i))$ 时 T 为 Hermite 变换;

(b) 当且仅当对所有 i 与 j 都有 $(T(u_j), u_i) = -(u_j, T(u_i))$ 时 T 为斜 Hermite 变换.

证明 任取 V 中两个元素 x 与 y 并将它们表示成基元素的线性组合, 设 $x = \sum\limits_{j=1}^{n} x_j u_j$, $y = \sum\limits_{i=1}^{n} y_i u_i$, 则我们有

$$(T(x), y) = \left(\sum_{j=1}^{n} x_j T(u_j), y \right) = \sum_{j=1}^{n} x_j \left(T(u_j), \sum_{i=1}^{n} y_i u_i \right) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \overline{y_i} (T(u_j), u_i).$$

同理可得

$$(x, T(y)) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \overline{y_i} (u_j, T(u_i)).$$

由上述二式即得命题 (a) 与 (b). □

下面我们用 T 的矩阵表示来刻划这些性质.

定理 7.6 令 $(u_1,\cdots,u_n)$ 为 V 的一组标准正交基, $\boldsymbol{A}=(a_{ij})$ 是线性变换 $T:V\longrightarrow V$ 在这组基下的矩阵, 则我们有

(a) 当且仅当对所有 i 与 j 都有 $a_{ij}=\overline{a_{ji}}$ 时 T 是 Hermite 变换;

(b) 当且仅当对所有 i 与 j 都有 $a_{ij}=-\overline{a_{ji}}$ 时 T 是斜 Hermite 变换.

证明 由于 $\boldsymbol{A}$ 是线性变换 T 在基 $(u_1,\cdots,u_n)$ 下的矩阵, 因此我们有 $T(u_j)=\sum\limits_{k=1}^{n}a_{kj}u_k$. 取 $T(u_j)$ 与 u_i 的内积, 由内积的线性性得

$$(T(u_j),u_i)=\left(\sum_{k=1}^{n}a_{kj}u_k,u_i\right)=\sum_{k=1}^{n}a_{kj}(u_k,u_i).$$

由于当 $k\neq i$ 时都有 $(u_k,u_i)=0$ 而 $(u_i,u_i)=1$, 因此上面等式中的最后一个和式简化为 $a_{ij}(u_i,u_i)=a_{ij}$. 于是我们得

$$a_{ij}=(T(u_j),u_i),\quad \text{对所有}\, i,j.$$

交换上式中 i 与 j 的次序并取共轭, 再由内积的 Hermite 对称性即得

$$\overline{a}_{ji}=(u_j,T(u_i)),\quad \text{对所有}\, i,j.$$

于是由定理 7.5 即得结论. □

7.7 Hermite 矩阵和斜 Hermite 矩阵 · 伴随矩阵

由定理 7.6 我们引入下述定义.

定义 设 $\boldsymbol{A}=(a_{ij})$ 为一个方阵, 如果对所有 i 与 j 都有 $a_{ij}=\overline{a}_{ji}$, 则称 $\boldsymbol{A}$ 为一个**Hermite 矩阵**; 如果对所有 i 与 j 都有 $a_{ij}=-\overline{a}_{ji}$, 则称 $\boldsymbol{A}$ 为**斜 Hermite 矩阵**.

定理 7.6 是说, 若 T 为有限维空间 V 上的一个线性变换, 则 T 为 Hermite 变换 (或斜 Hermite 变换) 的充分必要条件是 T 在 V 的标准正交基下的矩阵为 Hermite 矩阵 (或斜 Hermite 矩阵).

这两类矩阵还可以用另外一种方式来刻划. 令 $\overline{\boldsymbol{A}}$ 表示将 $\boldsymbol{A}$ 中各元素用它的共轭复数代替后所得到的矩阵, 叫作 $\boldsymbol{A}$ 的**共轭矩阵** (conjugate matrix), 则 $\boldsymbol{A}$ 为 Hermite 矩阵的充分必要条件是 $\boldsymbol{A}$ 与其共轭转置矩阵 $\overline{\boldsymbol{A}}^{\mathrm{T}}$ 相等, 即 $\boldsymbol{A}=\overline{\boldsymbol{A}}^{\mathrm{T}}$; $\boldsymbol{A}$ 为斜 Hermite 矩阵的充分必要条件是 $\boldsymbol{A}=-\overline{\boldsymbol{A}}^{\mathrm{T}}$. 对共轭转置矩阵我们给它一个特殊的名字.

定义 (伴随矩阵) 任意矩阵 $\boldsymbol{A}$ 的共轭转置矩阵 $\overline{\boldsymbol{A}}^{\mathrm{T}}$ 叫作 $\boldsymbol{A}$ 的**伴随矩阵**, 记作 $\boldsymbol{A}^*$.

根据上述定义, 对一个方阵 $\boldsymbol{A}$, 若 $\boldsymbol{A}=\boldsymbol{A}^*$, 则 $\boldsymbol{A}$ 是 Hermite 矩阵; 若 $\boldsymbol{A}=-\boldsymbol{A}^*$, 则 $\boldsymbol{A}$ 是斜 Hermite 矩阵. 由于一个 Hermite 矩阵与它的伴随矩阵相等, 因此也叫

作自伴矩阵 (self-adjoint matrix).

注 早期关于矩阵的很多文献用"**伴随矩阵**"(adjoint matrix) 表示一个矩阵的**余子式矩阵** (cofactor matrix) 的转置矩阵, 这是一个完全不同的对象, 这里给出的定义与当前算子理论中的命名法相一致.

7.8 Hermite 矩阵与斜 Hermite 矩阵的对角化

定理 7.7 任一 $n\times n$ 的 Hermite 矩阵或斜 Hermite 矩阵 $\boldsymbol{A}$ 都相似于一个以 $\boldsymbol{A}$ 的特征值为主对角线元素的对角矩阵 $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$. 而且存在非奇异矩阵 $\boldsymbol{C}$ 使

$$\Lambda=\boldsymbol{C}^{-1}\boldsymbol{AC}.$$

此处 $\boldsymbol{C}$ 满足条件 $\boldsymbol{C}^{-1}=\boldsymbol{C}^*$, 即 $\boldsymbol{C}$ 的逆矩阵等于 $\boldsymbol{C}$ 的伴随矩阵.

证明 令 V 表示由复 n 序组构成的空间 $\mathbb{C}^n$, 再令 $(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n)$ 为由单位坐标向量组成的标准正交基. 对 $\boldsymbol{x}=\sum\limits_{i=1}^n x_i\boldsymbol{e}_i$ 与 $\boldsymbol{y}=\sum\limits_{i=1}^n y_i\boldsymbol{e}_i$, 由 $(\boldsymbol{x},\boldsymbol{y})=\sum\limits_{i=1}^n x_i\overline{y_i}$ 定义 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积. 对给定的矩阵 $\boldsymbol{A}$, 令 T 为在上面选定的基下以 $\boldsymbol{A}$ 为矩阵的线性变换. 由定理 7.4 可知, V 有由 T 的特征向量组成的标准正交基 $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$, T 在这组基下的矩阵为对角矩阵

$$\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n),$$

此处 λ_k 为对应于特征向量 $\boldsymbol{u}_k$ 的特征值. 由于 $\boldsymbol{A}$ 与 Λ 都是 T 的矩阵表示, 因此它们相似. 从而我们有 $\Lambda=\boldsymbol{C}^{-1}\boldsymbol{AC}$, 其中 $\boldsymbol{C}=(c_{ij})$ 是从基 $(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n)$ 到基 $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$ 的非奇异转移矩阵:

$$[\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n]=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]\boldsymbol{C}.$$

上述等式表明, $\boldsymbol{C}$ 的第 j 列由 $\boldsymbol{u}_j$ 在基 $(\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n)$ 下的坐标组成, 从而 c_{ij} 是 $\boldsymbol{u}_j$ 的第 i 个分量. $\boldsymbol{u}_j$ 与 $\boldsymbol{u}_i$ 的内积由

$$(\boldsymbol{u}_j,\boldsymbol{u}_i)=\sum_{k=1}^n c_{kj}\bar{c}_{ki}$$

给出. 由于 $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$ 为一组标准正交基, 因此 $\boldsymbol{CC}^*=\boldsymbol{I}$, 从而 $\boldsymbol{C}^{-1}=\boldsymbol{C}^*$. □

注 定理 7.7 的证明同时也给出了求对角化矩阵 $\boldsymbol{C}$ 的方法. 找出由特征向量组成的一组标准正交基 $(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_n)$, 然后将 $\boldsymbol{u}_j$ 在由单位坐标向量组成的基下的各个分量作为矩阵 $\boldsymbol{C}$ 的第 j 列的元素.

例 1 实 Hermite 矩阵

$$\boldsymbol{A}=\begin{bmatrix}2&2\\2&5\end{bmatrix}$$

有特征值 $\lambda_1=1$ 和 $\lambda_2=6$. 属于 1 的特征向量为 $t(2,-1)^{\mathrm{T}},t\neq 0$; 属于 6 的特征向量为 $t(1,2)^{\mathrm{T}},t\neq 0$. 取 $t=1/\sqrt{5}$, 则特征向量 $\boldsymbol{u}_1=t(2,-1)^{\mathrm{T}}$ 与 $\boldsymbol{u}_2=t(1,2)^{\mathrm{T}}$ 组成一个标准正交组, 从而

$$C = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

是 A 的一个对角化矩阵. 由于 C 为实矩阵, 因而 $C^* = C^{\mathrm{T}}$. 不难验证下式成立:

$$C^{\mathrm{T}}AC = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}. \qquad \square$$

例 2 如果 A 本来就已经是对角矩阵了, 则定理 7.7 中的对角化矩阵 C 或者保持 A 不变, 或者仅仅改变 A 中主对角线元素的次序. □

7.9 酉矩阵 · 正交矩阵

定义 设 A 为一个方阵, 若 $AA^* = I$, 则称 A 为一个**酉矩阵** (unitary); 若 $AA^{\mathrm{T}} = I$, 则 A 称为**正交矩阵** (orthogonal matrix).

注 每一个实酉矩阵都是正交矩阵, 这是因为此时有 $A^* = A^{\mathrm{T}}$.

我们乐于使用酉矩阵和正交矩阵, 这是因为这两种矩阵的求逆十分容易, 只需取共轭和转置即可. 酉矩阵的逆矩阵是它的伴随矩阵, 正交矩阵的逆矩阵则是它的转置矩阵.

如果 A 是酉矩阵, 则 A 的行列式是一个绝对值为 1 的复数. 如果 A 是正交矩阵, 则 A 的行列式为 ± 1. 这是因为 $\det A = \det A^{\mathrm{T}}$ 及 $\det \overline{A} = \overline{(\det A)}$, 所以如果 A 为酉矩阵则有 $\det(AA^*) = \det I = 1$, 而 $\det(AA^*) = (\det A)(\det A^*) = (\det A)(\det \overline{A}) = |\det A|^2$, 因此 $\det A$ 是绝对值为 1 的一个复数. 当 A 为正交矩阵时, 同理可得 $(\det A)^2 = 1$, 由于 $\det A$ 为实数, 因此 $\det A = \pm 1$.

注 由酉矩阵和正交矩阵表示的线性变换将在 7.19 节中讨论.

定理 7.7 告诉我们, Hermite 矩阵与斜 Hermite 矩阵总可以用一个酉矩阵将其对角化. 实 Hermite 矩阵 (即实对称矩阵) 的特征值都是实数而且其特征向量也都可以取实向量, 因此我们有下述结论.

定理 7.7 的推论 任意一个实 Hermite 矩阵总可以用实正交矩阵将其对角化. □

上述命题对实斜 Hermite 矩阵并不成立 (见 7.10 节习题 11).

我们再给出下列相关概念.

定义 设 A 为一个实方阵或复方阵, 若 $A = A^{\mathrm{T}}$, 则称 A 为**对称矩阵**; 若 $A = -A^{\mathrm{T}}$, 则称 A 为**斜对称矩阵**.

例 1 设 $A = \begin{bmatrix} 1+i & 2 \\ 3-i & 4i \end{bmatrix}$, 则

$$\overline{A} = \begin{bmatrix} 1-i & 2 \\ 3+i & -4i \end{bmatrix}, \qquad A^{\mathrm{T}} = \begin{bmatrix} 1+i & 3-i \\ 2 & 4i \end{bmatrix}, \qquad A^* = \begin{bmatrix} 1-i & 3+i \\ 2 & -4i \end{bmatrix}. \qquad \square$$

例 2 如果 $\boldsymbol{A}$ 是实矩阵, 则 $\boldsymbol{A}$ 的伴随矩阵就是它的转置矩阵, 即 $\boldsymbol{A}^* = \boldsymbol{A}^{\mathrm{T}}$. 因此每一个实 Hermite 矩阵都是对称矩阵. 然而, 对称矩阵并不都是 Hermite 矩阵. □

例 3 $\begin{bmatrix}1 & 2\\2 & 3\end{bmatrix}$ 与 $\begin{bmatrix}1 & 2+i\\2-i & 3\end{bmatrix}$ 都是 Hermite 矩阵, 但前者是对称矩阵, 而后者不是对称矩阵. □

例 4 $\begin{bmatrix}0 & -2\\2 & 0\end{bmatrix}$ 与 $\begin{bmatrix}i & -2\\2 & 3i\end{bmatrix}$ 都是斜 Hermite 矩阵, 但前者是斜对称矩阵而后者则不是. □

例 5 Hermite 矩阵的主对角线元素都是实数. 斜 Hermite 矩阵的主对角线元素都是纯虚数. 斜对称矩阵的主对角线元素都是零. □

例 6 对任意方阵 $\boldsymbol{A}$, 矩阵 $\boldsymbol{B} = \frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^*)$ 都是 Hermite 矩阵, 而矩阵 $\boldsymbol{C} = \frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^*)$ 则是斜 Hermite 矩阵且 $\boldsymbol{B}+\boldsymbol{C}=\boldsymbol{A}$. 因此每一个方阵 $\boldsymbol{A}$ 都可以表示成 $\boldsymbol{A}=\boldsymbol{B}+\boldsymbol{C}$ 的形式, 其中 $\boldsymbol{B}$ 是 Hermite 矩阵, $\boldsymbol{C}$ 是斜 Hermite 矩阵. 验证这种分解的唯一性是很简单的. 同样, 每一个方阵 $\boldsymbol{A}$ 都可以唯一地表示成一个对称矩阵与一个斜对称矩阵之和: $\boldsymbol{A} = \frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}) + \frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}})$. □

7.10 习 题

1. 在下列各矩阵中确定何者为对称矩阵, 何者为斜对称矩阵, 何者为 Hermite 矩阵及何者为斜 Hermite 矩阵.

   (a) $\begin{bmatrix}0 & 1 & 2\\1 & 0 & 3\\2 & 3 & 4\end{bmatrix}$. (b) $\begin{bmatrix}0 & i & 2\\i & 0 & 3\\-2 & -3 & 4i\end{bmatrix}$. (c) $\begin{bmatrix}0 & i & 2\\-i & 0 & 3\\-2 & -3 & 0\end{bmatrix}$. (d) $\begin{bmatrix}0 & 1 & 2\\-1 & 0 & 3\\-2 & -3 & 0\end{bmatrix}$.

2. (a) 验证 2×2 矩阵 $\boldsymbol{A} = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$ 是一个正交矩阵.

   (b) 设 T 是在通常的基 $(\boldsymbol{i},\boldsymbol{j})$ 下的矩阵为上述矩阵 $\boldsymbol{A}$ 的线性变换, 证明: T 将平面上极坐标为 (r,α) 的点映为极坐标为 $(r,\alpha+\theta)$ 的点, 即 T 把平面绕原点转动 θ 角度.

3. 设 $V=\mathbb{R}^3$ 并且取通常的基向量 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$. 证明: 下列各小题中的矩阵都是正交矩阵并且是各题中所指定的变换的矩阵表示.

   (a) $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{bmatrix}$ (关于 xy 平面的反射变换). (b) $\begin{bmatrix}1 & 0 & 0\\0 & -1 & 0\\0 & 0 & -1\end{bmatrix}$ (关于 x 轴的反射变换).

   (c) $\begin{bmatrix}-1 & 0 & 0\\0 & -1 & 0\\0 & 0 & -1\end{bmatrix}$ (关于原点的反射变换). (d) $\begin{bmatrix}1 & 0 & 0\\0 & \cos\theta & -\sin\theta\\0 & \sin\theta & \cos\theta\end{bmatrix}$ (绕 x 轴的旋转).

   (e) $\begin{bmatrix}-1 & 0 & 0\\0 & \cos\theta & -\sin\theta\\0 & \sin\theta & \cos\theta\end{bmatrix}$ (关于 x 轴旋转之后再作关于 yz 平面的反射变换)

4. 设 $\boldsymbol{A}$ 为实正交矩阵, 若 $\det \boldsymbol{A} = 1$, 则 $\boldsymbol{A}$ 叫正常 (proper)**正交矩阵**; 若 $\det \boldsymbol{A} = -1$, 则称 $\boldsymbol{A}$ 为非正常 (improper) **正交矩阵**.

(a) 若 $\boldsymbol{A}$ 是 2×2 正常正交矩阵, 证明存在某个 θ 使 $\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 矩阵 $\boldsymbol{A}$ 表示转动 θ 角的一个旋转变换.

(b) 证明 $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 与 $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ 都是非正常正交矩阵. 前一矩阵表示 xy 平面上关于 x 轴的反射变换; 后一矩阵表示关于 y 轴的反射变换, 找出所有 2×2 非正常正交矩阵.

在习题 5~8 中, (a) 求 $\boldsymbol{A}$ 的一个正交特征向量组; (b) 求酉矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{-1}\boldsymbol{AC}$ 为对角矩阵.

5. $\boldsymbol{A} = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$,　　6. $\boldsymbol{A} = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$,

7. $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$,　　8. $\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$.

9. 确定下列各矩阵中哪些是酉矩阵, 哪些是正交矩阵 (a, b, θ 都是实数).

(a) $\begin{bmatrix} e^{ia} & 0 \\ 0 & e^{ib} \end{bmatrix}$.　(b) $\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$.　(c) $\begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{3}\sqrt{3} & \frac{1}{6}\sqrt{6} \\ 0 & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} & \frac{1}{3}\sqrt{3} & -\frac{1}{6}\sqrt{6} \end{bmatrix}$.

10. 狭义相对论用到下列方程:

$$x' = a(x - vt),\quad y' = y,\quad z' = z,\quad t' = a(t - vx/c^2).$$

此处 v 表示运动物体的速度, c 是光速, $a = c/\sqrt{c^2 - v^2}$. 将 (x, y, z, t) 映为 (x', y', z', t') 的线性变换叫**Lorentz 变换**.

(a) 令 $(x_1, x_2, x_3, x_4) = (x, y, z, ict)$ 及 $(x_1', x_2', x_3', x_4') = (x', y', z', ict')$, 证明: 上面 4 个方程可表示为下面的矩阵方程:

$$[x_1', x_2', x_3', x_4'] = [x_1, x_2, x_3, x_4]\begin{bmatrix} a & 0 & 0 & -iav/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ iav/c & 0 & 0 & a \end{bmatrix}.$$

(b) 证明: (a) 中的 4×4 矩阵是正交矩阵但不是酉矩阵.

11. 设 a 为非零实数. $\boldsymbol{A} = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$ 为斜对称方阵.

(a) 求 $\boldsymbol{A}$ 的一个标准正交特征向量组.

(b) 求酉矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{-1}\boldsymbol{AC}$ 为对角矩阵.

(c) 证明: 不存在实正交矩阵 $\boldsymbol{C}$ 使 $\boldsymbol{C}^{-1}\boldsymbol{AC}$ 为对角矩阵.

12. 证明: 如果一个 Hermite 矩阵或斜 Hermite 矩阵 $\boldsymbol{A}$ 的所有特征值都等于 c, 则 $\boldsymbol{A} = c\boldsymbol{I}$.

13. 证明: 如果 $\boldsymbol{A}$ 是实斜对称矩阵, 则 $\boldsymbol{I} - \boldsymbol{A}$ 与 $\boldsymbol{I} + \boldsymbol{A}$ 都是非奇异矩阵并且 $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1}$ 是正交矩阵.

14. 证明下列关于 $n\times n$ 矩阵的命题, 或者给出一个反例.

(a) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是酉矩阵, 则 $\boldsymbol{A} + \boldsymbol{B}$ 也是酉矩阵.

(b) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是酉矩阵, 则 $\boldsymbol{A} + \boldsymbol{B}$ 不是酉矩阵.

(c) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 都是酉矩阵, 则 $\boldsymbol{AB}$ 也是酉矩阵.

(d) 如果 $\boldsymbol{A}$ 与 $\boldsymbol{AB}$ 都是酉矩阵, 则 $\boldsymbol{B}$ 也是酉矩阵.

7.11　二　次　型

现在我们转向关于二次型的研究, 这是实对称变换最重要的应用之一.

设 V 为实欧氏空间, $T: V \longrightarrow V$ 为一个对称算子, 即在内积中 T 可以从作用在一个因子上转到作用在另一个因子上:

$$(T(x), y) = (x, T(y)), \qquad \text{对 } V \text{ 中所有的 } x \text{ 与 } y.$$

我们取定 T, 再由

$$Q(x) = (T(x), x)$$

在 V 上定义一个实值函数 Q, 叫作与 T 相伴的**二次型** (quadratic form). 下述定理指出, 在有限维的情形, $Q(x)$ 是 x 的各分量的一个二次多项式, 因此我们将 $Q(x)$ 叫作与 T 相伴的二次型.

定理 7.8　设 $(u_1, \cdots, u_n)$ 是实欧氏空间 V 的一组标准正交基, 再设 $T: V \longrightarrow V$ 为对称线性变换, $\boldsymbol{A} = (a_{ij})$ 为 T 在基 $(u_1, \cdots, u_n)$ 下的矩阵, 则二次型 $Q(x) = (T(x), x)$ 与 $\boldsymbol{A}$ 之间有如下关系:

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad \text{若 } x = \sum_{i=1}^{n} x_i u_i. \tag{7.3}$$

证明　由线性性, 我们有 $T(x) = \sum_{i=1}^{n} x_i T(u_i)$, 从而得

$$Q(x) = \left(\sum_{i=1}^{n} x_i T(u_i), \sum_{j=1}^{n} x_j u_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j (T(u_i), u_j).$$

由定理 7.6, 我们有 $a_{ij} = a_{ji} = (T(u_i), u_j)$, 于是即得关系式 (7.3). □

即使矩阵 $\boldsymbol{A}$ 不是对称矩阵, (7.3) 中的和式也是有意义的.

定义　设 V 为任意一个实欧氏空间, $(u_1, \cdots, u_n)$ 为它的一组标准正交基, 再令 $\boldsymbol{A} = (a_{ij})$ 为任意一个 $n \times n$ 实矩阵. 对 V 中任意点 $x = \sum_{i=1}^{n} x_i u_i$, 由公式

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{7.4}$$

所定义的实值函数 Q 叫作与 $\boldsymbol{A}$ 相伴的**二次型**.

如果 $\boldsymbol{A}$ 是对角矩阵, 则当 $i \neq j$ 时都有 $a_{ij} = 0$, 因此 (7.4) 中的和式只包含平方项并且可表示成下述简单形式:

$$Q(x) = \sum_{i=1}^{n} a_{ii} x_i^2,$$

这时的二次型叫作**对角二次型** (diagonal form).

公式 (7.4) 中的 2 重和也可以表示为三个矩阵的乘积.

定理 7.9 设 $\boldsymbol{X}=[x_1,\cdots,x_n]$ 为 $1\times n$ 行矩阵, $\boldsymbol{A}=(a_{ij})$ 为 $n\times n$ 矩阵, 则 $\boldsymbol{XAX}^{\mathrm{T}}$ 为下述 1×1 矩阵:

$$\boldsymbol{XAX}^{\mathrm{T}}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j. \tag{7.5}$$

证明 矩阵乘积 $\boldsymbol{XA}$ 是一个 $1\times n$ 矩阵, 即 $\boldsymbol{XA}=[y_1,\cdots,y_n]$, 其中的元素 y_j 是 $\boldsymbol{X}$ 与 $\boldsymbol{A}$ 的第 j 个列向量的点积:

$$y_j=\sum_{i=1}^{n}x_ia_{ij}.$$

因此乘积 $\boldsymbol{XAX}^{\mathrm{T}}$ 是一个 1×1 矩阵, 它仅有的一个元素是 $\boldsymbol{XA}$ 与 $\boldsymbol{X}^{\mathrm{T}}$ 的点积:

$$\sum_{j=1}^{n}y_jx_j=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}x_ia_{ij}\right)x_j=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j. \qquad \square$$

注 依惯例, 我们将 1×1 矩阵 $\boldsymbol{XAX}^{\mathrm{T}}$ 就看作和式 (7.5) 并将矩阵乘积 $\boldsymbol{XAX}^{\mathrm{T}}$ 叫作一个二次型. 公式 (7.4) 可以简单地写作.

$$Q(x)=\boldsymbol{XAX}^{\mathrm{T}}.$$

例 1 设 $\boldsymbol{X}=[x_1,x_2]$, $\boldsymbol{A}=\begin{bmatrix}1&-1\\-3&5\end{bmatrix}$, 则我们有

$$\boldsymbol{XA}=[x_1,x_2]\begin{bmatrix}1&-1\\-3&5\end{bmatrix}=[x_1-3x_2,-x_1+5x_2],$$

因此得

$$\boldsymbol{XAX}^{\mathrm{T}}=[x_1-3x_2,-x_1+5x_2]\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1^2-3x_2x_1-x_1x_2+5x_2^2. \qquad \square$$

例 2 设 $\boldsymbol{X}=[x_1,x_2]$, $\boldsymbol{B}=\begin{bmatrix}1&-2\\-2&5\end{bmatrix}$, 则

$$\boldsymbol{XBX}^{\mathrm{T}}=[x_1,x_2]\begin{bmatrix}1&-2\\-2&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1^2-2x_2x_1-2x_1x_2+5x_2^2. \qquad \square$$

在上述两例中, 将两个混合积项相加都得到 $-4x_1x_2$, 因此 $\boldsymbol{XAX}^{\mathrm{T}}=\boldsymbol{XBX}^{\mathrm{T}}$. 这两个例子显示, 不同的矩阵可能会导致相同的二次型. 注意到, 其中的一个矩阵, 即矩阵 $\boldsymbol{B}$, 是对称矩阵. 受上述讨论的启发, 我们有下述定理.

定理 7.10 对任意 $n\times n$ 矩阵 $\boldsymbol{A}$ 和任意 $1\times n$ 行矩阵 $\boldsymbol{X}$, 都有 $\boldsymbol{XAX}^{\mathrm{T}}=\boldsymbol{XBX}^{\mathrm{T}}$, 其中 $\boldsymbol{B}=\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}})$ 为对称矩阵.

证明 由于 $\boldsymbol{XAX}^{\mathrm{T}}$ 为 1×1 矩阵, 因此它等于自己的转置矩阵, 即 $\boldsymbol{XAX}^{\mathrm{T}}=(\boldsymbol{XAX}^{\mathrm{T}})^{\mathrm{T}}$. 由于矩阵乘积的转置矩阵等于各矩阵的转置矩阵按相反顺序的乘积 (见 4.21 节习题 7(d)), 因此我们有 $(\boldsymbol{XAX}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{XA}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}$, 从而 $\boldsymbol{XAX}^{\mathrm{T}}=\frac{1}{2}\boldsymbol{XAX}^{\mathrm{T}}+\frac{1}{2}\boldsymbol{XA}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{XBX}^{\mathrm{T}}$. $\square$

7.12 将实二次型化为对角形

实对称矩阵 $\boldsymbol{A}$ 是一个 Hermite 矩阵, 因此, 由定理 7.7, $\boldsymbol{A}$ 相似于对角矩阵 $\Lambda=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, 其中 $\lambda_1,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值, 而且由定理 7.7 的推论知, 存在正交矩阵 $\boldsymbol{C}$ 使得 $\Lambda=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$. 下面我们证明, $\boldsymbol{C}$ 可以用来将二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 化为对角形.

定理 7.11 设 $\boldsymbol{XAX}^{\mathrm{T}}$ 是与实对称矩阵 $\boldsymbol{A}$ 相伴的二次型, 再设 $\boldsymbol{C}$ 是将 $\boldsymbol{A}$ 化为对角矩阵的正交矩阵, 即 $\Lambda=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC}$, 则我们有

$$\boldsymbol{XAX}^{\mathrm{T}}=\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}=\sum_{i=1}^{n}\lambda_i y_i^2,$$

其中 $\boldsymbol{Y}=[y_1,\cdots,y_n]$ 为行矩阵 $\boldsymbol{Y}=\boldsymbol{XC}$, $\lambda_1,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的特征值.

证明 由于 $\boldsymbol{C}$ 是正交矩阵, 因此 $\boldsymbol{C}^{-1}=\boldsymbol{C}^{\mathrm{T}}$, 从而由 $\boldsymbol{Y}=\boldsymbol{XC}$ 得 $\boldsymbol{X}=\boldsymbol{YC}^{\mathrm{T}}$, 于是得

$$\boldsymbol{XAX}^{\mathrm{T}}=(\boldsymbol{YC}^{\mathrm{T}})\boldsymbol{A}(\boldsymbol{YC}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{Y}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{AC})\boldsymbol{Y}^{\mathrm{T}}=\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}. \qquad \square$$

注 定理 7.11 是说, 线性变换 $\boldsymbol{Y}=\boldsymbol{XC}$ 将二次型 $\boldsymbol{XAX}^{\mathrm{T}}$ 化为对角形 $\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}$.

例 1 与单位矩阵相伴的二次型是

$$\boldsymbol{XIX}^{\mathrm{T}}=\sum_{i=1}^{n}x_i^2=\|\boldsymbol{X}\|^2,$$

即它是向量 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 的长度的平方. 设 $\boldsymbol{C}$ 为一个正交矩阵, 则线性变换 $\boldsymbol{Y}=\boldsymbol{XC}$ 给出一个新的二次型 $\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}$, 这里 $\Lambda=\boldsymbol{CIC}^{\mathrm{T}}=\boldsymbol{CC}^{\mathrm{T}}=\boldsymbol{I}$. 因为 $\boldsymbol{XIX}^{\mathrm{T}}=\boldsymbol{YIY}^{\mathrm{T}}$, 因此我们得 $\|\boldsymbol{X}\|^2=\|\boldsymbol{Y}\|^2$, 从而 $\boldsymbol{Y}$ 与 $\boldsymbol{X}$ 的长度相同. 保持每一个向量的长度都不变的线性变换叫作**等距变换** (isometry). 这类变换将在 7.19 节中作详细讨论. $\square$

例 2 求将二次型

$$Q(x)=2x_1^2+4x_1x_2+5x_2^2$$

化为对角形的正交矩阵 $\boldsymbol{C}$.

解 我们将混合积项写成 $2x_1x_2+2x_2x_1$, 令 $\boldsymbol{A}=\begin{bmatrix}2&2\\2&5\end{bmatrix}$, 则得 $Q(x)=\boldsymbol{XAX}^{\mathrm{T}}$. 对称矩阵 $\boldsymbol{A}$ 已经在定理 7.7 后面的例 1 中被对角化. $\boldsymbol{A}$ 的特征值为 $\lambda_1=1$, $\lambda_2=6$, 并且特征向量 $\boldsymbol{u}_1=\frac{1}{\sqrt5}(2,-1)$ 与 $\boldsymbol{u}_2=\frac{1}{\sqrt5}(1,2)$ 组成一组标准正交基, 因此得到一个正交的对角化矩阵 $\boldsymbol{C}=\frac{1}{\sqrt5}\begin{bmatrix}2&1\\-1&2\end{bmatrix}$, 相应的对角形二次型为

$$Y\Lambda Y^{\mathrm{T}}=\lambda_1y_1^2+\lambda_2y_2^2=y_1^2+6y_2^2. \qquad \square$$

例 2 的结果有一个简单的几何解释, 如图 7.1 所示. 线性变换 $\boldsymbol{Y}=\boldsymbol{XC}$ 可以看作是将基 $\boldsymbol{i},\boldsymbol{j}$ 映到一组新基 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 上的一个旋转变换. 在第一组基下坐标为 (x_1,x_2) 的点在第二组基下的坐标是 (y_1,y_2). 由于 $\boldsymbol{XAX}^{\mathrm{T}}=\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}$, 因此对某个 c, 满足方程 $\boldsymbol{XAX}^{\mathrm{T}}=c$ 的点 (x_1,x_2) 的集合和满足方程 $\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}=c$ 的点 (y_1,y_2) 的集合是一样的. 后一个方程可以写作 $y_1^2+6y_2^2=c$, 若 $c>0$, 则它就是某一个椭圆的笛卡儿方程. 因此方程 $\boldsymbol{XAX}^{\mathrm{T}}=c$, 即 $2x_1^2+4x_1x_2+5x_2^2=c$, 表示在原来坐标系中的是同一个椭圆. 图 7.1 表示当 $c=9$ 时的椭圆.

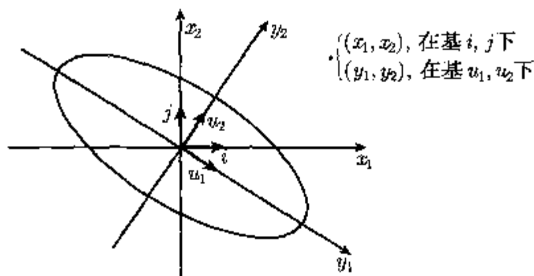


图 7.1 由一个正交矩阵给出的坐标轴的旋转. 在 x_1x_2 坐标系中, 椭圆的笛卡儿方程为 $\boldsymbol{XAX}^{\mathrm{T}}=9$, 而在 y_1y_2 坐标系中椭圆的笛卡儿方程为 $\boldsymbol{Y}\Lambda\boldsymbol{Y}^{\mathrm{T}}=9$.

7.13 对二次曲线的应用

利用化二次型为对角形可以确定平面上满足形如

$$ax^2+bxy+cy^2+dx+ey+f=0 \tag{7.6}$$

的笛卡儿方程的所有点 (x,y) 的集合, 我们将会发现这样的集合都是**二次曲线**, 即椭圆、双曲线、抛物线, 或下面的退化情形之一 (空集、单独一点、一条或两条直线). 二次曲线的类型由方程的二次项所决定, 也就是说由二次型 $ax^2+bxy+cy^2$ 所决定. 为了和前面几节中所用的记号一致, 我们将 x 记作 x_1, 将 y 记作 x_2, 并把此二次型表示成如下矩阵乘积的形式:

$$\boldsymbol{XAX}^{\mathrm{T}}=ax_1^2+bx_1x_2+cx_2^2,$$

其中 $\boldsymbol{X}=[x_1,x_2]$ 而

$$\boldsymbol{A}=\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$$

是一个对称矩阵. 由旋转变换 $\boldsymbol{Y}=\boldsymbol{XC}$ (或者旋转变换之后再关于某个坐标轴作反射变换) 可将此二次型化为对角形 $\lambda_1y_1^2+\lambda_2y_2^2$, 此处 λ_1 与 λ_2 是 $\boldsymbol{A}$ 的特征值. 一组标准正交的特征向量 $\boldsymbol{u}_1,\boldsymbol{u}_2$ 确定一组新的坐标系, 在这组新坐标系下, 笛卡儿方程 (7.6) 化为

$$\lambda_1y_1^2+\lambda_2y_2^2+d'y_1+e'y_2+f=0, \tag{7.7}$$

其中 d' 与 e' 是新的线性项的系数. 在这个方程中没有混合积项 y_1y_2, 因此二次曲线的类型可以很容易地通过特征值 λ_1 和 λ_2 来确定.

如果二次曲线是非退化的, 当 λ_1 与 λ_2 正负号相同时方程 (7.7) 表示一个**椭圆** (ellipse); 当 λ_1 与 λ_2 正负号相反时, 则表示一组**双曲线** (hyperbola); 当 λ_1 与 λ_2 中有一个等于零时, 则表示一条**抛物线** (parabola). 这三种情形分别对应于 $\lambda_1\lambda_2 > 0$, $\lambda_1\lambda_2 < 0$ 与 $\lambda_1\lambda_2 = 0$. 我们分别举例说明之.

例 1 $2x^2 + 4xy + 5y^2 + 4x + 13y - \frac{1}{4} = 0$, 我们将这个方程重写为

$$2x_1^2 + 4x_1x_2 + 5x_2^2 + 4x_1 + 13x_2 - \frac{1}{4} = 0. \tag{7.8}$$

我们在上一节例 2 中已经讨论过二次型 $2x_1^2+4x_1x_2+5x_2^2$, 它的矩阵有特征值 $\lambda_1 = 1$ 与 $\lambda_2 = 6$, 并且 $\boldsymbol{u}_1 = \frac{1}{\sqrt{5}}(2, -1)$ 与 $\boldsymbol{u}_2 = \frac{1}{\sqrt{5}}(1, 2)$ 是它的一组标准正交的特征向量, 因此 $\boldsymbol{C} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ 是一个正交对角化矩阵. 利用 $\boldsymbol{C}$ 可将方程 (7.8) 中的二次型化为 $y_1^2 + 6y_2^2$. 为确定这一变换对线性部分的影响, 我们将旋转变换的方程 $\boldsymbol{Y} = \boldsymbol{XC}$ 改写成 $\boldsymbol{X} = \boldsymbol{YC}^{\mathrm{T}}$, 于是得

$$[x_1, x_2] = \frac{1}{\sqrt{5}}[y_1, y_2]\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \quad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

由此, 线性部分 $4x_1 + 13x_2$ 变为

$$\frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{13}{\sqrt{5}}(-y_1 + 2y_2) = -\sqrt{5}\,y_1 + 6\sqrt{5}\,y_2.$$

变换后的笛卡儿方程为

$$y_1^2 + 6y_2^2 - \sqrt{5}\,y_1 + 6\sqrt{5}\,y_2 - \frac{1}{4} = 0.$$

再分别对 y_1 与 y_2 用配方法, 则上式可改写为如下形式:

$$\left(y_1 - \frac{1}{2}\sqrt{5}\right)^2 + 6\left(y_2 + \frac{1}{2}\sqrt{5}\right)^2 = 9.$$

这是在 y_1y_2 坐标系中以点 $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ 为中心的一个椭圆的方程. 坐标轴 y_1 与 y_2 的正向分别由特征向量 $\boldsymbol{u}_1$ 与 $\boldsymbol{u}_2$ 所决定, 如图 7.2 所示.

若令

$$z_1 = y_1 - \frac{1}{2}\sqrt{5}, \qquad z_2 = y_2 + \frac{1}{2}\sqrt{5},$$

则椭圆的方程还可进一步简化. 从几何上说, 这相应于引入一个新的坐标系, 它的两条坐标轴分别与 y_1 和 y_2 平行, 但取椭圆的中心作为坐标原点. 将新坐标系记作 z_1z_2 坐标系, 则在新坐标系中, 椭圆的方程简化为

$$z_1^2 + 6z_2^2 = 9 \quad 或 \quad \frac{z_1^2}{9} + \frac{z_2^2}{3/2} = 1.$$

这个椭圆及所有三个坐标系都在图 7.2 中给出. □

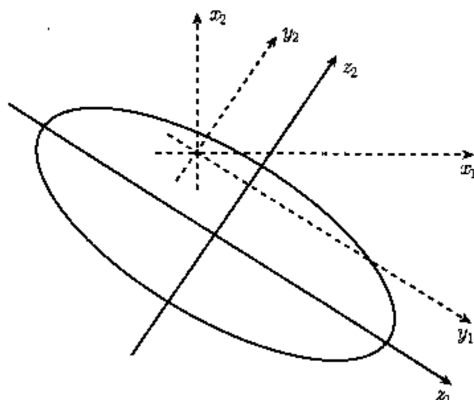


图 7.2 坐标轴的旋转与平移. 在旋转变换 $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{C}$ 之后再作平移变换 $z_1 = y_1 - \frac{1}{2}\sqrt{5}$, $z_2 = y_2 + \frac{1}{2}\sqrt{5}$

例 2 $2x^2 - 4xy - y^2 - 4x + 10y - 13 = 0$. 将此方程重写为

$$2x_1^2 - 4x_1x_2 - x_2^2 - 4x_1 + 10x_2 - 13 = 0,$$

它的二次型部分是 $\boldsymbol{X}\boldsymbol{A}\boldsymbol{X}^{\mathrm{T}}$, 此处 $\boldsymbol{A} = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}$. 矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = 3$ 与 $\lambda_2 = -2$. 令 $\boldsymbol{u}_1 = \frac{1}{\sqrt{5}}(2, -1)$, $\boldsymbol{u}_2 = \frac{1}{\sqrt{5}}(1, 2)$, 则 $\boldsymbol{u}_1$ 与 $\boldsymbol{u}_2$ 是标准正交特征向量组, 于是 $\boldsymbol{C} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ 是一个正交的对角化矩阵. 由旋转变换的方程 $\boldsymbol{X} = \boldsymbol{Y}\boldsymbol{C}^{\mathrm{T}}$ 得

$$x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \quad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

从而经变换后的椭圆方程变为

$$3y_1^2 - 2y_2^2 - \frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{10}{\sqrt{5}}(-y_1 + 2y_2) - 13 = 0,$$

经整理后得

$$3y_1^2 - 2y_2^2 - \frac{18}{\sqrt{5}}y_1 + \frac{16}{\sqrt{5}}y_2 - 13 = 0.$$

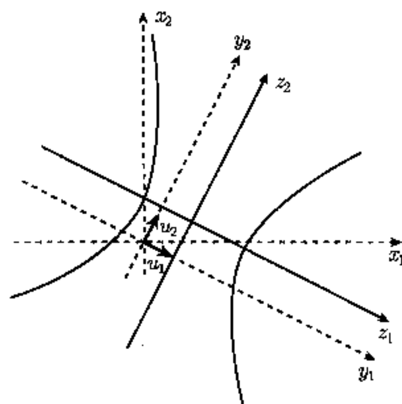
分别对 y_1 与 y_2 配方, 则得方程

$$3\left(y_1 - \frac{3}{5}\sqrt{5}\right)^2 - 2\left(y_2 - \frac{4}{5}\sqrt{5}\right)^2 = 12,$$

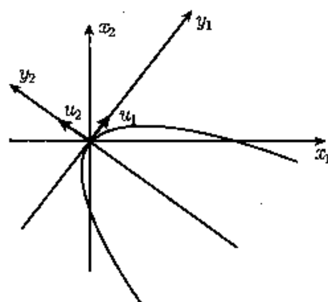
它代表在 y_1y_2 平面上以点 $(\frac{3}{5}\sqrt{5},\frac{4}{5}\sqrt{5})$ 为中心的双曲线. 再作平移变换 $z_1=y_1-\frac{3}{5}\sqrt{5}$, $z_2=y_2-\frac{4}{5}\sqrt{5}$. 则上述方程可进一步简化为

$$3z_1^2-2z_2^2=12 \quad 或 \quad \frac{z_1^2}{4}-\frac{z_2^2}{6}=1.$$

此双曲线的图像见图 7.3(a). 特征向量 $\boldsymbol{u}_1$ 与 $\boldsymbol{u}_2$ 决定了坐标轴 y_1 与 y_2 的正向. □



(a) 双曲线 $3z_1^2-2z_2^2=12$



(b) 抛物线 $y_1^2+y_2=0$

图 7.3　例 2 与例 3 中的曲线

例 3　$9x^2+24xy+16y^2-20x+15y=0$. 将此方程改写为

$$9x_1^2+24x_1x_2+16x_2^2-20x_1+15x_2=0,$$

它的二次型部分的矩阵为 $\boldsymbol{A}=\begin{bmatrix}9&12\\12&16\end{bmatrix}$, $\boldsymbol{A}$ 的特征值为 $\lambda_1=25$, $\lambda_2=0$. 令 $\boldsymbol{u}_1=\frac{1}{5}(3,4)$, $\boldsymbol{u}_2=\frac{1}{5}(-4,3)$, 则 u_1 与 u_2 组成一个标准正交特征向量组, 于是 $\boldsymbol{C}=\frac{1}{5}\begin{bmatrix}3&-4\\4&3\end{bmatrix}$ 是一个正交对角化矩阵. 由旋转变换的方程 $\boldsymbol{X}=\boldsymbol{Y}\boldsymbol{C}^{\mathrm{T}}$ 得

$$x_1=\frac{1}{5}(3y_1-4y_2), \quad x_2=\frac{1}{5}(4y_1+3y_2).$$

经变换后的笛卡儿方程为

$$25y_1^2-\frac{20}{5}(3y_1-4y_2)+\frac{15}{5}(4y_1+3y_2)=0,$$

整理之, 得 $y_1^2+y_2=0$, 这是一条以坐标原点为顶点的抛物线, 其图形见图 7.3(b).□

例 4 (退化情形)　此时仅仅由特征值还不足以判断此笛卡儿方程代表的是哪一类退化的二次曲线. 例如, 下面三个方程

$$x^2+2y^2=1, \quad x^2+2y^2=0, \quad x^2+2y^2=-1,$$

它们有相同的特征值, 但第一个方程代表的是一个非退化的椭圆, 第二个方程仅表示一点 $(x,y)=(0,0)$, 第三个方程表示的是空集. 后面两种情形可看作椭圆的退化

情形.

方程 $y^2=0$ 的图像是 x 轴, 方程 $y^2-1=0$ 表示一对平行直线 $y=1$ 与 $y=-1$, 它们可看作抛物线的退化情形; 方程 $x^2-4y^2=0$ 代表两条相交直线, 因为由 $x^2-4y^2=0$ 得 $x-2y=0$ 或 $x+2y=0$, 这可以看作双曲线的退化情形.

然而, 如果笛卡儿方程 $ax^2+bxy+cy^2+dx+ey+f=0$ 代表的是非退化二次曲线, 则曲线的**类型** (type) 可以相当容易地确定. 二次型 $ax^2+bxy+cy^2$ 的矩阵的特征多项式为

$$\det\begin{bmatrix}\lambda-a & -b/2\\ -b/2 & \lambda-c\end{bmatrix}=\lambda^2-(a+c)\lambda+\left(ac-\frac{1}{4}b^2\right)=(\lambda-\lambda_1)(\lambda-\lambda_2).$$

从而两个特征值的乘积等于常数项, 即

$$\lambda_1\lambda_2=ac-\frac{1}{4}b^2=\frac{1}{4}(4ac-b^2).$$

由于二次曲线的类型是由乘积 $\lambda_1\lambda_2$ 为正、为负或为零所决定的, 因此当 $4ac-b^2$ 为正、为负或为零时, 二次曲线分别为椭圆、双曲线或抛物线. $4ac-b^2$ 叫作二次型 $ax^2+bxy+cy^2$ 的**判别式** (discriminant), 它就是我们在处理定理 2.19 中的二次曲线时所遇到的判别式, 不过那里和判别式相关的是二次曲线的离心率, 而这里和判别式相关的是二次型的矩阵表示的特征值的乘积. 在例 1~3 中, 判别式的值分别是 34, −24 和 0. □

7.14 习 题

在习题 1~7 中, (a) 求二次型的对称矩阵 $\boldsymbol{A}$; (b) 求 $\boldsymbol{A}$ 的特征值; (c) 求一个标准正交特征向量组; (d) 求正交对角化矩阵 $\boldsymbol{C}$.

1. $4x_1^2+4x_1x_2+x_2^2$.
2. x_1x_2.
3. $x_1^2+2x_1x_2-x_2^2$.
4. $34x_1^2-24x_1x_2+41x_2^2$.
5. $x_1^2+x_1x_2+x_1x_3+x_2x_3$.
6. $2x_1^2+4x_1x_3+x_2^2-x_3^2$.
7. $3x_1^2+4x_1x_2+8x_1x_3+4x_2x_3+3x_3^2$.

在习题 8~18 中, 求所给笛卡儿方程表示的二次曲线并画出其草图.

8. $y^2-2xy+2x^2-5=0$.
9. $y^2-2xy+5x=0$.
10. $y^2-2xy+x^2-5x=0$.
11. $5x^2-4xy+2y^2-6=0$.
12. $19x^2+4xy+16y^2-212x+104y=356$.
13. $9x^2+24xy+16y^2-52x+14y=6$.
14. $5x^2+6xy+5y^2-2=0$.
15. $x^2+2xy+y^2-2x+2y+3=0$.

16. $2x^2+4xy+5y^2-2x-y-4=0$.
17. $x^2+4xy-2y^2-12=0$.
18. $xy+y-2x-2=0$.
19. 当 c 取什么值时笛卡儿方程 $2xy-4x+7y+c=0$ 的图像是一对直线?
20. 由长半轴为 A 与短半轴为 B 的椭圆所围成的图形的面积是 πAB. 设方程 $ax^2+bxy+cy^2=1$ 所代表的二次曲线是一个椭圆, 证明由此椭圆所围成的区域的面积是 $2\pi/\sqrt{4ac-b^2}$, 这个式子给出了判别式 $4ac-b^2$ 的一个几何意义.

7.15 正定二次型

作为定理 7.11 的另外一个应用, 我们要建立二次型的正定性或负定性与它的特征值之间的联系.

定理 7.12 设 V 为任意实欧氏空间, $(u_1,\cdots,u_n)$ 为 V 的一组标准正交基, 令 $\boldsymbol{A}=(a_{ij})$ 为一个 $n\times n$ 实对称矩阵, 再令

$$Q(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j,$$

其中 $x=\sum\limits_{i=1}^{n}x_iu_i$, 则

(a) 对所有 $x\neq O$ 都有 $Q(x)>0$ 的充分必要条件是 $\boldsymbol{A}$ 的所有特征值都大于零;

(b) 对所有 $x\neq O$ 都有 $Q(x)<0$ 的充分必要条件是 $\boldsymbol{A}$ 的所有特征值都小于零.

注 在情形 (a) 中的二次型叫正定 (positive definite) 二次型; 情形 (b) 中的二次型叫负定 (negative definite) 二次型.

证明 对 $x=\sum\limits_{i=1}^{n}x_iu_i$, 作行矩阵 $\boldsymbol{X}=[x_1,\cdots,x_n]$. 由定理 7.11, 存在一个正交矩阵 $\boldsymbol{C}$ 将二次型 $Q(x)=\boldsymbol{XAX}^{\mathrm{T}}$ 化为对角形:

$$Q(x)=\sum_{i=1}^{n}\lambda_iy_i^2,$$

此处 $\boldsymbol{Y}=[y_1,\cdots,y_n]$ 是行矩阵 $\boldsymbol{Y}=\boldsymbol{XC}$, $\lambda_1,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的特征值. 由于 $\boldsymbol{A}$ 是实对称矩阵, 因而其特征值都是实数.

若所有特征值都是正数, 则当 $\boldsymbol{Y}\neq\boldsymbol{O}$ 时, 对角形二次型中的和式也是正数. 而由于 $\boldsymbol{Y}=\boldsymbol{XC}$, 因此 $\boldsymbol{X}=\boldsymbol{YC}^{-1}$, 从而当且仅当 $x\neq O$ 时有 $\boldsymbol{Y}\neq\boldsymbol{O}$, 于是对所有 $x\neq O$ 都有 $Q(x)>0$, 即得充分性.

反之, 若对所有 $x\neq O$ 都有 $Q(x)>0$, 我们可选取 x 使 $\boldsymbol{Y}=\boldsymbol{XC}$ 是第 k 个基向量 $\boldsymbol{u}_k$. 对这个 x, 对角形二次型 $Q(x)$ 只包含一项, 即 $Q(x)=\lambda_k$, 因此每个 λ_k 都大于零, 即得必要性. 这就证明了 (a). 同理可证 (b). □

*7.16[1] 由二次型的值求对称变换的特征值

现在我们去掉空间 V 为有限维的要求, 并给出对称变换的特征值与这个变换的二次型的值之间的联系.

设 x 是属于特征值 λ 的一个范数为 1 的特征向量, 则由 $T(x)=\lambda x$ 及 $(x,x)=1$, 我们有

$$Q(x)=(T(x),x)=(\lambda x,x)=\lambda(x,x)=\lambda. \tag{7.9}$$

V 中所有满足 $(x,x)=1$ 的元素 x 的集合叫作 V 中的**单位球面** (unit sphere). 等式 (7.9) 证明了下述定理.

定理 7.13 设 $T:V\longrightarrow V$ 为实欧氏空间 V 上的一个对称线性变换, 令 $Q(x)=(T(x),x)$, 若 T 有特征值, 则其特征值必可由 $Q(x)$ 在 V 的单位球面上取得. □

例 1 令 $V=\mathbb{R}^2$, 取 $(\boldsymbol{i},\boldsymbol{j})$ 为基, 并取通常的点积作为内积. 设 T 是以 $\boldsymbol{A}=\begin{bmatrix}4&0\\0&8\end{bmatrix}$ 为矩阵的对称线性变换. 则 T 的二次型是由

$$Q(x)=\sum_{i=1}^{2}\sum_{j=1}^{2}a_{ij}x_ix_j=4x_1^2+8x_2^2$$

给出的对角形二次型. T 的特征值为 $\lambda_1=4$ 和 $\lambda_2=8$. 易知这两个特征值分别是 Q 在单位圆 $x_1^2+x_2^2=1$ 上的最小值与最大值. 事实上, 在单位圆上, 我们有

$$Q(x)=4(x_1^2+x_2^2)+4x_2^2=4+4x_2^2,\qquad -1\leqslant x_2\leqslant 1.$$

因此当 $x_2=0$ 时 $Q(x)$ 取最小值 4, 当 $x_2=\pm1$ 时, $Q(x)$ 取最大值 8.

图 7.4 给出了单位圆和两个椭圆. 里面的椭圆的笛卡儿方程为 $4x_1^2+8x_2^2=4$, 它由平面上所有满足条件 $Q(x)=4$ 的点 $x=(x_1,x_2)$ 组成. 外面的椭圆的笛卡儿方程为 $4x_1^2+8x_2^2=8$, 它由所有满足条件 $Q(x)=8$ 的点组成, 里面椭圆与单位圆的切点 $(\pm1,0)$ 是属于特征值 $\lambda_1=4$ 的特征向量, 外面椭圆上的点 $(0,\pm1)$ 是属于特征值 $\lambda_2=8$ 的特征向量. □

上面各例说明了在一般情形也成立的特征值的极值性质. 在下一节中, 我们将证明: 若最小特征值与最大特征值存在, 则它们必定是 Q 在单位球上取得的最小值与最大值. 关于这些极值性质的讨论要用到关于二次型的下述定理. 应该指出的是, 这个定理并不要求空间 V 是有限维的.

定理 7.14 设 $T:V\longrightarrow V$ 为实欧氏空间 V 上的一个对称线性变换, T 的二次型为 $Q(x)=(T(x),x)$. 假定 Q 在 V 上不改变正负号, 则当对 V 中某个 x 有

[1] 加星号的章节可以略去不读或以后再读, 这不影响连续性.

$Q(x)=0$ 时也有 $T(x)=O$. 换言之, 如果 Q 不改变正负号, 则只有当 x 属于 T 的零空间时才有 $Q(x)=0$.

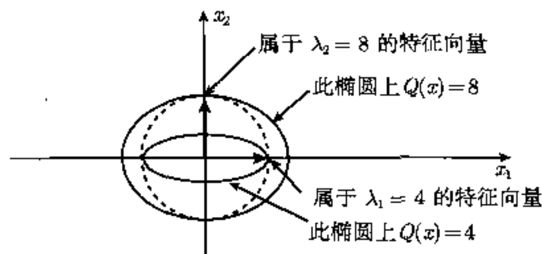


图 7.4　二维情形下, T 的特征值与 Q 在单位球上的值之间的几何关系

证明　假设对 V 中某个 x 有 $Q(x)=0$, 令 y 为 V 中任意元素. 任选实数 t 并将 $Q(x+ty)$ 看作 t 的函数, 由 T 的线性性、内积的线性性以及 T 的对称性, 我们有

$$\begin{aligned}Q(x+ty)&=(T(x+ty),\,x+ty)=(T(x)+tT(y),\,x+ty)\\&=(T(x),\,x)+t(T(x),\,y)+t(T(y),\,x)+t^2(T(y),\,y)\\&=Q(x)+2t(T(x),\,y)+t^2Q(y)=0+2at+bt^2,\end{aligned}$$

其中 $a=(T(x),\,y)$, $b=Q(y)$. 如果 Q 在 V 上非负, 则得下述不等式:

$$2at+bt^2\geqslant 0,\qquad \text{对所有实数 } t.$$

我们要证明由此必有 $a=0$. 若 $b=0$, 则除非 $a=0$, 否则上述不等式不可能对所有 t 都成立. 若 $b>0$, 则二次多项式 $p(t)=2at+bt^2$ 的图像是一条开口向上的抛物线并且与 t 轴不可能有多余 1 个的交点. 由于 $p(0)=0$, 我们一定有 $a=0$, 若不然, 则 $p(t)$ 将有另一个根 $t=-b/(2a)$, 而这是不可能的. 由于 $a=(T(x),\,y)$, 从而由 $a=0$ 即得 $(T(x),\,y)=0$ 对所有 y 都成立. 特别地, 当 $y=T(x)$ 时得 $(T(x),\,T(x))=0$, 从而 $T(x)=O$.

若 Q 在 V 上是非正的, 则对所有 t 都有 $p(t)=2at+bt^2\leqslant 0$. 对 $-p(t)$ 作上述论证即可得 $a=0$. □

*7.17　对称线性变换的最大值与最小值

现在我们来证明二次型在单位球面上的极值就是特征值.

定理 7.15　设 $T:V\longrightarrow V$ 为实欧氏空间 V 上的一个对称线性变换, 其二次型为 $Q(x)=(T(x),\,x)$. 假定 Q 在单位球面上存在最大值与最小值[①], Q 可能是最

① 若 V 是无限维空间, 则二次型 Q 未必在单位球面上取最大值与最小值, 当 T 没有特征值时这种情形就会发生. 若 V 是有限维空间, Q 总会在单位球面上的某处取得最大值或最小值, 这可由关于连续函数最大值与最小值的一个更一般的定理的推论而得.

大值也可能是最小值, 并且 Q 在满足 $(u, u) = 1$ 的点 u 处取得最值, 则 u 是 T 的一个特征向量, 与它相应的特征值是 Q 在单位球面上的最值 $Q(u)$.

证明 假定 Q 在点 u 处取最小值, 则

$$Q(x) \geqslant Q(u), \qquad \text{对所有满足条件 } (x,x) = 1 \text{ 的 } x. \tag{7.10}$$

令 $\lambda = Q(u)$, 若 $(x, x) = 1$, 则 $Q(u) = \lambda(x, x) = (\lambda x, x)$, 因此不等式 (7.10) 可以表示成

$$(T(x), x) \geqslant (\lambda x, x), \qquad \text{对所有满足条件 } (x, x) = 1 \text{ 的 } x. \tag{7.11}$$

我们证明不等式 (7.11) 对 V 中所有的 x 都成立. 设 $\|x\| = a$, 则有 $x = ay$, 其中 $\|y\| = 1$, 因此得

$$(T(x), x) = (T(ay), ay) = a^2(T(y), y) \quad \text{与} \quad (\lambda x, x) = (\lambda ay, ay) = a^2(\lambda y, y).$$

而由 $(y, y) = 1$ 及不等式 (7.11) 我们有 $(T(y), y) \geqslant (\lambda y, y)$, 将此不等式两边同乘以 a^2 即可知不等式 (7.11) 对 $x = ay$ 也成立.

由 $(T(x), x) - (\lambda x, x) = (T(x) - \lambda x, x)$, 我们可将不等式 (7.11) 表示成以下形式:

$$(S(x), x) \geqslant 0, \quad \text{此处} \quad S = T - \lambda I. \tag{7.12}$$

当 $x = u$ 时不等式 (7.10) 中等号成立, 于是不等式 (7.12) 中等号也成立. 由于线性变换 S 是对称的, 不等式 (7.12) 是说由 $Q_1(x) = (S(x), x)$ 给出的二次型 Q_1 在 V 上非负. 由于当 $x = u$ 时我们有 $Q_1(u) = 0$, 因此, 由定理 7.14 必有 $S(u) = 0$, 即 $T(u) = \lambda u$, 因此 u 是 T 的一个特征向量, 而 $\lambda = Q(u)$ 是对应于 u 的特征值. 由此完成了当 Q 在 u 处取得最小值时定理的证明.

如果在 u 处取最大值, 将上述证明中所有的不等式反向, 将定理 7.14 应用于非正二次型 Q_1 即得结论. □

*7.18 有限维情形

今设 $\dim V = n$, 则 T 有 n 个实特征值, 将它们依递增顺序排列:

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n.$$

由定理 7.15, 最小特征值 λ_1 是 Q 在单位球面上的最小值, 最大特征值 λ_n 是 Q 在单位球面上的最大值. 我们将进一步指出中间的那些特征值也是 Q 在单位球面的某些子集上的最大值或最小值.

设 u_1 是单位球面上的一个特征向量, 使得 Q 在 u_1 处取得最小值, 则 $\lambda_1 = Q(u_1)$. 设 λ 为一个不同于 λ_1 的特征值, 则任一属于 λ 的特征向量必与 u_1 正交, 我们自然希望在由 u_1 生成的子空间 S 的正交补空间中去寻求这样的一个特征向量.

S 的正交补空间 $S^{\perp}$ 由 V 中所有与 u_1 正交的元素组成. 特别地, 对特征值 $\lambda \neq \lambda_1$, $S^{\perp}$ 包含所有属于 λ 的特征向量. 不难验证 $\dim S^{\perp} = n-1$ 并且 T 将 $S^{\perp}$ 映入 $S^{\perp}$ 自身.[①] 令 S_{n-1} 表示 $n-1$ 维子空间 $S^{\perp}$ 中的单位球面 (S_{n-1} 是 V 中单位球面的一个子集). 将定理 7.15 用于子空间 $S^{\perp}$, 我们得 $\lambda_2 = Q(u_2)$, 其中 u_2 是 S_{n-1} 上使 Q 取得最小值的一点.

同理, 令 S_{n-2} 为由所有与 u_1 和 u_2 都正交的元素组成的 $n-2$ 维空间中的单位球面, 则特征值 λ_3 等于 Q 在单位球面 S_{n-2} 上的极小值. 如此继续下去, 可知第 k 个特征值 λ_k 是 Q 在一个 $n-k+1$ 维子空间的单位球面 S_{n-k+1} 上的最小值. 这些最小值中最大的那一个, 即 λ_n, 也是 Q 在各个单位球面 S_{n-k+1} 上取得的最大值. 特征向量 $u_1, \cdots, u_n$ 组成 V 的一组标准正交基.

7.19　酉　变　换

现在我们转向另一类重要的线性变换——酉变换. 在有限维的情形, 酉变换由酉矩阵表示. 在本节中我们并不要求所论及的空间的维数有限.

定义　设 E 为欧氏空间, V 为 E 的子空间, $T: V \longrightarrow E$ 为线性空间. 如果有

$$(T(x), T(y)) = (x, y), \quad \text{对 } V \text{ 中所有元素 } x \text{ 与 } y, \tag{7.13}$$

则称 T 为 V 上的**酉变换** (unitary transformation). 如果 E 是实欧氏空间, 酉变换也叫作**正交变换**.

等式 (7.13) 是说 T 是保持内积不变的线性变换. 人们自然希望 T 也保持正交性和向量的范数不变, 因为它们都是由内积导出的.

定理 7.16　设 $T: V \longrightarrow E$ 是 V 上的一个酉变换, 则对 V 中所有的 x 与 y 都有下列结论.

(a) 由 $(x, y) = 0$ 得 $(T(x), T(y)) = 0$　(T 保持正交性).

(b) $\|T(x)\| = \|x\|$　(T 保持范数不变).

(c) $\|T(x) - T(y)\| = \|x - y\|$　(T 保持距离不变).

(d) T 可逆并且 T^{-1} 是 $T(V)$ 上的酉变换.

证明　由等式 (7.13) 即得 (a). 在等式 (7.13) 中令 $y = x$ 即得 (b). 由于 $T(x) - T(y) = T(x-y)$, 因此由 (b) 即得 (c).

最后, 我们利用 (b) 证明 (d). 由 (b) 可知, 若 $T(x) = O$ 则 $x = O$, 因此 T 可逆. 如果 $x \in T(V)$ 及 $y \in T(V)$, 则存在 $u, v \in V$ 使得 $x = T(u)$, $y = T(v)$, 于是我们有

$$(T^{-1}(x), T^{-1}(y)) = (u, v) = (T(u), T(v)) = (x, y),$$

---

① 这一点已在 7.5 节定理 7.4 的证明中给出.

因此 T^{-1} 满足等式 (7.13), 从而 T^{-1} 是 $T(V)$ 上的酉变换. □

关于酉变换的特征值与特征向量, 我们有下述定理.

定理 7.17 设 $T: V \longrightarrow E$ 为 V 上的一个酉变换:

(a) 若 T 有特征值 λ, 则 $|\lambda| = 1$.

(b) 属于不同特征值的特征向量正交.

(c) 若 $V = E$, $\dim V = n$, 并且 V 为复欧氏空间, 则存在 T 的 n 个特征向量 $u_1, \cdots, u_n$, 它们组成 V 的一组标准正交基, T 在这组基下的矩阵是对角矩阵 $\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$. 其中 λ_k 是对应于特征向量 u_k 的特征值.

证明 为证 (a), 令 x 为属于 λ 的一个特征向量, 则 $x \neq O$ 且 $T(x) = \lambda x$. 在等式 (7.13) 中取 $y = x$ 则得

$$(\lambda x, \lambda x) = (x, x) \qquad \text{即 } \lambda\overline{\lambda}(x, x) = (x, x).$$

由 $(x, x) > 0$ 及 $\lambda\overline{\lambda} = |\lambda|^2$ 即得 $|\lambda| = 1$.

为证 (b), 令 x 与 y 分别为属于不同特征值 λ 与 μ 的特征向量, 因此有 $T(x) = \lambda x$, $T(y) = \mu y$, 此处 $\lambda \neq \mu$. 我们用两种方法计算内积 $(T(x), T(y))$. 由于 T 是酉变换, 因此我们有

$$(T(x), T(y)) = (x, y).$$

又因 x 与 y 是特征向量, 因此又有

$$(T(x), T(y)) = (\lambda x, \mu y) = \lambda\overline{\mu}(x, y).$$

从而 $\lambda\overline{\mu}(x, y) = (x, y)$, 因此若 $\lambda\overline{\mu} \neq 1$ 则必有 $(x, y) = 0$. 而由 (a) 有 $\lambda\overline{\lambda} = 1$, 因此若 $\lambda\overline{\mu} = 1$ 则应有 $\lambda\overline{\lambda} = \lambda\overline{\mu}$, 从而 $\overline{\lambda} = \overline{\mu}$. 这与 $\lambda \neq \mu$ 相矛盾. 因此 $\lambda\overline{\mu} \neq 1$, 从而 $(x, y) = 0$.

最后我们对 n 作归纳法来证明 (c), 这里所用的方法和关于 Hermite 变换的相应结果的定理 7.4 的证法几乎完全一样, 唯一不同的是关于 T 将 $S^{\perp}$ 映入它自身的证明, 此处

$$S^{\perp} = \{x : x \in V \text{ 且 } (x, u_1) = 0\},$$

其中 u_1 是 T 的属于特征值 λ_1 的特征向量. 由等式 $T(u_1) = \lambda_1 u_1$, $\lambda_1\overline{\lambda}_1 = |\lambda_1|^2 = 1$ 得

$$u_1 = \lambda_1^{-1} T(u_1) = \overline{\lambda}_1 T(u_1).$$

任取 $x \in S^{\perp}$, 注意到

$$\big(T(x), u_1\big) = \big(T(x), \overline{\lambda}_1 T(u_1)\big) = \lambda_1\big(T(x), T(u_1)\big) = \lambda_1(x, u_1) = 0.$$

因此若 $x \in S^{\perp}$ 则 $T(x) \in S^{\perp}$, 从而 T 将 $S^{\perp}$ 映入 $S^{\perp}$. 证明的其余部分与定理 7.4 的完全一样, 这里就不重复了. □

下面两个定理刻划了有限维空间中酉变换的性质, 我们这里只给出证明的要点.

定理 7.18　设 $\dim V = n$, 令 $U = (u_1, \cdots, u_n)$ 为 V 中取定的一组基, 则线性变换 $T : V \longrightarrow V$ 为酉变换的充分必要条件是

$$\big(T(u_i), T(u_j)\big) = (u_i, u_j), \quad 对所有\ i\ 与\ j. \tag{7.14}$$

特别地, 若 U 是标准正交基, 则 T 为酉变换的充分必要条件是 T 把 U 映为标准正交基.

证明要点　令 $x = \sum\limits_{i=1}^{n} x_i u_i$, $y = \sum\limits_{j=1}^{n} y_j u_j$, 则我们有

$$(x, y) = \left(\sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \overline{y_j}(u_i, u_j),$$

$$(T(x), T(y)) = \left(\sum_{i=1}^{n} x_i T(u_i), \sum_{j=1}^{n} y_j T(u_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \overline{y}_j (T(u_i), T(u_j)),$$

比较 (x, y) 与 $(T(x), T(y))$ 即可得结论. □

定理 7.19　设 $\dim V = n$, 令 $(u_1, \cdots, u_n)$ 为 V 的一组标准正交基, 令 $\boldsymbol{A} = (a_{ij})$ 为线性变换 $T : V \longrightarrow V$ 在这组基下的矩阵, 则 T 为酉变换的充分必要条件是 $\boldsymbol{A}$ 为酉矩阵, 即

$$\boldsymbol{A}^* \boldsymbol{A} = \boldsymbol{I}. \tag{7.15}$$

证明要点　由于 (u_i, u_j) 等于单位矩阵 $\boldsymbol{I}$ 的 (i,j)- 元, 因此由等式 (7.15) 得

$$(u_i, u_j) = \sum_{k=1}^{n} \overline{a}_{ki} a_{kj} = \sum_{k=1}^{n} a_{ki} \overline{a}_{kj}. \tag{7.16}$$

又由于 $\boldsymbol{A}$ 是 T 在给定基下的矩阵, 因此 $T(u_i) = \sum\limits_{k=1}^{n} a_{ki} u_k$, $T(u_j) = \sum\limits_{r=1}^{n} a_{rj} u_r$, 从而得

$$\begin{aligned}(T(u_i), T(u_j)) &= \left(\sum_{k=1}^{n} a_{ki} u_k, \sum_{r=1}^{n} a_{rj} u_r\right) = \sum_{k=1}^{n}\sum_{r=1}^{n} a_{ki} \overline{a}_{rj} (u_k, u_r) \\ &= \sum_{k=1}^{n} a_{ki} \overline{a}_{kj}.\end{aligned}$$

将上式与等式 (7.16) 相比较, 由定理 7.18 即可得结论. □

定理 7.20　每一个酉矩阵 $\boldsymbol{A}$ 都具有下述性质.

(a) $\boldsymbol{A}$ 为非奇异矩阵并且 $\boldsymbol{A}^{-1} = \boldsymbol{A}^*$.

(b) $\boldsymbol{A}^{\mathrm{T}}$, $\overline{\boldsymbol{A}}$ 与 $\boldsymbol{A}^*$ 都是酉矩阵.

(c) $\boldsymbol{A}$ 的特征值都是绝对值等于 1 的复数.

(d) $|\det \boldsymbol{A}| = 1$, 且当 $\boldsymbol{A}$ 为实矩阵时, $\det \boldsymbol{A} = \pm 1$. □

定理 7.20 的证明作为习题留给读者.

7.20 习 题

1. (a) 设 $T: V \longrightarrow V$ 是由 $T(x) = cx$ 定义的变换, 其中 c 为一给定的纯量. 证明: 当且仅当 $|c| = 1$ 时 T 为酉变换.
   (b) 设 V 是一维空间, 证明: 除了在 (a) 中给出的酉变换之外, V 上没有别的酉变换. 特别地, 若 V 是实一维空间, 则只有两个正交变换, 即 $T(x) = x$ 与 $T(x) = -x$.
2. 证明关于 $n \times n$ 实正交矩阵 $\boldsymbol{A}$ 的下列各命题.
   (a) 若 λ 是 $\boldsymbol{A}$ 的实特征值, 则 $\lambda = 1$ 或 $\lambda = -1$.
   (b) 若 λ 是 $\boldsymbol{A}$ 的一个复特征值, 则 λ 的共轭复数 $\overline{\lambda}$ 也是 $\boldsymbol{A}$ 的特征值. 换言之, $\boldsymbol{A}$ 的非实特征值必共轭成对出现.
   (c) 若 n 是奇数, 则 $\boldsymbol{A}$ 至少有一个实特征值.
3. 设 V 为 n 维实欧氏空间, 行列式为 1 的正交变换 $T: V \longrightarrow V$ 叫作一个**旋转变换** (rotation). 若 n 为奇数, 证明 1 是 T 的一个特征值. 这一结果说明奇数维空间的任一旋转变换都有一根固定轴. [提示: 利用习题 2 的结果.]
4. 给定一个以 -1 为 k 重特征值的实正交矩阵, 证明 $\det \boldsymbol{A} = (-1)^k$.
5. 证明: 如果 T 是一个**保持范数不变** (norm-preserving) 的线性变换, 则 T 是酉变换.
6. 证明: 如果 $T: V \to V$ 既是酉变换又是 Hermite 变换, 则 $T^2 = I$.
7. 给定 n 维欧氏空间的两组标准正交基, 证明: 存在一个酉变换 T 将其中一组基映为另一组基.
8. 求使下面的矩阵为酉矩阵的实数 a:

$$\begin{bmatrix} a & \frac{1}{2}i & \frac{1}{2}a(2i-1) \\ ia & \frac{1}{2}(1+i) & \frac{1}{2}a(1-i) \\ a & -\frac{1}{2} & \frac{1}{2}a(2-i) \end{bmatrix}.$$

9. 证明: 若 $\boldsymbol{A}$ 为斜 Hermite 矩阵, 则 $\boldsymbol{I} - \boldsymbol{A}$ 与 $\boldsymbol{I} + \boldsymbol{A}$ 都是非奇异矩阵并且乘积 $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1}$ 是酉矩阵.
10. 证明: 若 $\boldsymbol{A}$ 为酉矩阵且 $\boldsymbol{I} + \boldsymbol{A}$ 为非奇异矩阵, 则 $(\boldsymbol{I} - \boldsymbol{A})(\boldsymbol{I} + \boldsymbol{A})^{-1}$ 是斜 Hermite 矩阵.
11. 证明: 若 $\boldsymbol{A}$ 为 Hermite 矩阵, 则 $\boldsymbol{A} - i\boldsymbol{I}$ 为非奇异矩阵且 $(\boldsymbol{A} - i\boldsymbol{I})^{-1}(\boldsymbol{A} + i\boldsymbol{I})$ 为酉矩阵.
12. 证明: 任一酉矩阵都可以用酉矩阵对角化.
13. 一个矩阵 $\boldsymbol{A}$ 若满足条件 $\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A}$, 则称 $\boldsymbol{A}$ 为**正规矩阵** (normal matrix). 试确定以下各类型的矩阵中哪几类是正规矩阵.
    (a) Hermite 矩阵. (b) 斜 Hermite 矩阵. (c) 对称矩阵.
    (d) 斜对称矩阵. (e) 酉矩阵. (f) 正交矩阵.
14. 证明: 若 $\boldsymbol{A}$ 为正规矩阵 $(\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A})$, $\boldsymbol{U}$ 是酉矩阵, 则 $\boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U}$ 是正规矩阵.

*7.21 作用在函数空间上的对称变换和斜对称变换

在本章最后, 我们给出作用在实函数空间上的几个对称变换和斜对称变换的例子. 这些例子需要读者具有微积分的知识.

我们先来回忆 7.2 节中给出的定义: 设 E 为欧氏空间, V 是 E 的一个子空间, 若 V 上的一个线性变换 $T: V \to E$ 满足条件

$$(T(x), y) = (x, T(y)), \quad \text{对 } V \text{ 中所有的 } x \text{ 与 } y,$$

则称 T 是一个 Hermite 变换. 若内积是实数, 则称此 Hermite 变换为对称变换. 若

$$(T(x),y)=-(x,T(y)),\quad \text{对 } V \text{ 中所有的 } x \text{ 与 } y,$$

则称 T 是 V 上的一个斜 Hermite 变换. 当内积为实数时, 则称此斜 Hermite 变换为斜对称变换.

例 1 (空间 $C(a,b)$ 中的对称变换与斜对称变换) 令 $C(a,b)$ 表示闭区间 $[a,b]$ 上全体连续实函数组成的线性空间, 它的实内积为

$$(f,g)=\int_a^b f(t)g(t)\mathrm{d}t.$$

令 V 为 $C(a,b)$ 的一个子空间, 若 $T:V\to C(a,b)$ 是一个线性变换, 则

$$(f,T(g))=\int_a^b f(t)Tg(t)\mathrm{d}t,$$

这里我们用 $Tg(t)$ 表示 $T(g)(t)$. 在这一情形下, 关于对称性和斜对称性的要求变为

$$\int_a^b\{f(t)Tg(t)-g(t)Tf(t)\}\mathrm{d}t=0,\quad \text{若 } T \text{ 对称} \tag{7.17}$$

$$\int_a^b\{f(t)Tg(t)+g(t)Tf(t)\}\mathrm{d}t=0,\quad \text{若 } T \text{ 斜对称} \tag{7.18}$$

□

例 2 (与给定函数的乘积) 在例 1 的空间 $C(a,b)$ 中, 选取一个固定的函数 p 并定义 $T(f)$ 为 p 与 f 的积, 即 $T(f)=pf$. 因为被积函数为零, 所以 $C(a,b)$ 中所有的 f 与 g 都满足方程 (7.17). 因此, 与一个给定函数的乘积是一个对称变换. □

例 3 (微分变换) 在例 1 的空间 $C(a,b)$ 中, 令 V 为由所有在开区间 (a,b) 上有连续导数并且满足边界条件 $f(a)=f(b)$ 的函数 f 组成的子空间. 令 $D:V\to C(a,b)$ 为由 $D(f)=f'$ 给出的微分变换, 不难证明 D 是斜对称的. 在这一情形下, 公式 (7.18) 中的被积函数是乘积 fg 的导数, 因此积分等于

$$\int_a^b (fg)'(t)\mathrm{d}t=f(b)g(b)-f(a)g(a).$$

由于 f 与 g 都满足边界条件, 因此 $f(b)g(b)-f(a)g(a)=0$, 从而由边界条件可知 D 是斜对称变换. 在子空间 V 中仅有的特征函数是不为零的常数函数, 它们都属于特征值 0. □

例 4 (Sturm-Liouville 变换) 这个例子在线性二阶微分方程理论中很重要. 我们再次用到例 1 中的空间 $C(a,b)$. 令 V 为由在 $[a,b]$ 中有连续二阶导数并且满足下面两个边界条件

$$p(a)f(a)=0,\quad p(b)f(b)=0, \tag{7.19}$$

的所有函数 f 组成的子空间, 此处 p 是 $C(a,b)$ 中一个在 $[a,b]$ 上具有连续导数的固定函数. 令 q 为 $C(a,b)$ 中另一个固定函数并令 $T:V\to C(a,b)$ 为由方程

$$T(f)=(pf')'+qf$$

定义的变换, 我们称它为**Sturm-Liouville 变换**. 为验证它是对称变换, 注意到

$$fT(g) - gT(f) = f \cdot (pg')' - g \cdot (pf')'.$$

将上式代入公式 (7.17) 并对 $\int_a^b f \cdot (pg')'\mathrm{d}t$ 与 $\int_a^b g \cdot (pf')'\mathrm{d}t$ 同时施行分部积分. 由于 f 与 g 满足边界条件 (7.19), 因此我们有

$$\int_a^b \{f(t)Tg(t) - g(t)Tf(t)\}\mathrm{d}t = fpg'\Big|_a^b - \int_a^b pg'f'\mathrm{d}t - gpf'\Big|_a^b + \int_a^b pg'f'\mathrm{d}t = 0.$$

从而 T 是 V 上的对称变换. T 的特征函数是对某个实数 λ, 在 $[a,b]$ 上满足形如

$$(pf')' + qf = \lambda f$$

的微分方程及边界条件 (7.19) 的非零函数 f. □

例 5 定理 7.3 告诉我们, 对称变换的不同特征值对应的特征向量相互正交. 本例要将这一结果应用于那些在区间 $[a,b]$ 上满足形如

$$(pf')' + qf = \lambda f \tag{7.20}$$

的微分方程并且还满足边界条件 $p(a)f(a) = p(b)f(b) = 0$ 的非零函数 f. 所得到的结论是对应于 λ 的两个不同值的两个解 f 和 g 相互正交. 例如, 考察区间 $[0,\pi]$ 上的**简谐振动** 的微分方程

$$f'' + k^2 f = 0,$$

其中 $k \neq 0$. 这是微分方程 (7.20) 取 $p=1$, $q=0$ 及 $\lambda = -k^2$ 的情形, 所有的解由 $f(t) = c_1\cos kt + c_2\sin kt$ 给出. 由边界条件 $f(0)=0$ 得 $c_1 = 0$, 由另一个边界条件 $f(\pi)=0$ 得 $c_2\sin k\pi = 0$. 由于对非零解都有 $c_2 \neq 0$, 因此必有 $\sin k\pi = 0$, 即 k 为整数. 换言之, 当且仅当 k 为整数时存在满足边界条件的非零解, 这些解为 $f(t) = \sin nt$, $n = \pm1, \pm2, \cdots$. 由定理 7.3 给出的正交性条件即为下述我们熟知的关系式:

$$\int_0^\pi \sin nt \sin mt\mathrm{d}t = 0,$$

当 m^2 与 n^2 为不同的整数时上式成立. □

7.22 习 题

本节习题要求读者具有微积分知识

1. 令 $C(0,1)$ 为由在 $[0,1]$ 上连续的全体实函数组成的实线性空间, 并以 $(f,g) = \int_0^1 f(t)g(t)\mathrm{d}t$ 为内积. 令 V 为由 $C(0,1)$ 中所有使得 $\int_0^1 f(t)\mathrm{d}t = 0$ 的函数 f 组成的子空间, 并令 $T: V \to C(0,1)$ 为由 $Tf(x) = \int_0^x f(t)\mathrm{d}t$ 定义的积分变换. 证明: T 是斜对称变换.
2. 令 V 为由全体实多项式组成的欧氏空间且以 $(f,g) = \int_{-1}^1 f(t)g(t)\mathrm{d}t$ 为内积. 确定在下列各变换 $T: V \to V$ 中, 哪些是对称的, 哪些是斜对称的:
   (a) $Tf(x) = f(-x)$.
   (b) $Tf(x) = f(x)f(-x)$.
   (c) $Tf(x) = f(x) + f(-x)$.

(d) $Tf(x) = f(x) - f(-x)$.

3. 用内积

$$(f,g) = \int_a^b f(t)g(t)w(t)\mathrm{d}t,$$

和变换

$$T(f) = \frac{(pf')' + qf}{w}.$$

分别代替 7.21 节例 4 中的内积和 Sturm-Liouville 变换, 其中 w 是 $\boldsymbol{C}(a,b)$ 中一个固定的正函数. 证明: 此时这个变换是子空间 V 上的对称变换.

4. 本题表明 Legendre 多项式 (见 3.14 节) 是 Sturm-Liouville 变换的特征函数. Legendre 多项式由方程

$$P_n(t) = \frac{1}{2^n n!} f_n^{(n)}(t), \quad 其中\ f_n(t) = (t^2-1)^n.$$

定义.

(a) 验证 $(t^2-1)f_n'(t) = 2ntf_n(t)$.

(b) 对 (a) 中的方程微分 $n+1$ 次从而得到

$$(t^2-1)f_n^{(n+2)}(t)+2t(n+1)f_n^{(n+1)}(t)+n(n+1)f_n^{(n)}(t)=2ntf_n^{(n+1)}(t)+2n(n+1)f_n^{(n)}(t)$$

你可以利用乘积函数 $h(t) = f(t)g(t)$ 的 n 阶导数 $h^{(n)}(t)$ 的 Leibniz 公式:

$$h^{(n)}(t) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(t)g^{(n-k)}(t).$$

(c) 证明 (b) 中的方程可改写为如下形式:

$$[(t^2-1)P_n'(t)]' = n(n+1)P_n(t).$$

这表明 $P_n(t)$ 是由 $T(f) = (pf')'$ 所给出的区间 $[-1,1]$ 上的 Sturm-Liouville 变换 T 的特征函数, 此处 $p(t) = t^2-1$, $P_n(t)$ 是属于特征值 $\lambda = n(n+1)$ 的特征函数. 由于 $p(1) = p(-1) = 0$, 因此本例自动满足关于对称性的边界条件.

第 8 章　在线性微分方程中的应用

8.1　引　　言

线性代数最重要的应用之一就是在微分方程研究中的应用. 在物理世界, 当人们试图从已知的与变化率有关的知识来确定某些结论时, 就很自然地引出微分方程的问题. 例如, 我们可以从质点运动的速度 (位移的一阶导数) 或加速度 (位移的二阶导数) 设法求位移. 或者, 放射性物质以已知的速率裂变, 我们设法求在给定时刻物质的总量. 在这一类例子中, 问题是如何将一个未知函数 y 的导数 y' 和 y'' 中任意一个或两个所给出的信息表示成一个方程, 我们将这个方程称作**微分方程**(differential equation), 并由此求出未知函数 $y = f(x)$. 任意一个满足此微分方程的函数都叫作一个**解**(solution).

在一个微分方程中, 未知函数的最高阶导数的阶就叫作这个微分方程的**阶**(order). 根据这个定义, $y' = y$ 是一阶微分方程, 而 $y'' + 16y = 0$ 则是二阶微分方程. $y' = y$ 的所有解都可表示为 $y = ce^x$ 的形式, 其中 c 为任意常数. $y'' + 16y = 0$ 的所有解都可表示成 $y = c_1 \cos 4x + c_2 \sin 4x$ 的形式, 其中 c_1 与 c_2 为任意常数.

微分方程的历史始于 17 世纪, 那时 Newton、Leibnitz 与 Bernoullis 兄弟解出了从几何和力学问题提出的一些简单的一阶和二阶微分方程. 从这些早期的发现来看, 似乎从几何问题和物理问题提出的微分方程的解都可以用微积分学中熟知的初等函数来表示. 早期的许多工作的目标就是精心设计一些用初等方法解微分方程的技巧, 也就是用微积分学中熟悉的函数经有限次加法、减法、乘法、除法、复合以及求积等运算来解微分方程的技巧.

在 17 世纪末以前, 诸如分离变量法与积分因子法等微分方程的特殊解法的提出多少带有点偶然性. 到了 18 世纪, 人们才有较为系统的解法, 这类方法主要是由 Euler、Lagrange 与 Laplace 等人提出的. 人们很快就发现, 只有很少数的微分方程能用初等方法求解. 渐渐地, 数学家们开始意识到, 要想寻找能解所有微分方程的方法是毫无希望的. 然而, 他们发现, 研究给定的微分方程是否有解, 并且当它有解时如何从这个微分方程本身去推导出解的性质会更有成效. 在这样的框架下, 数学家们开始想到将微分方程看作是给出函数的新的源泉.

经验证明, 除去少数几类微分方程之外, 试图得到关于微分方程解的具有很大普遍性的结果是十分困难的. 在大量物理问题中提出的所谓**线性微分方程**, 就是这为数很少的几类微分方程中的一类. 下面几节将概括给出关于这类方程的某些主要结果.

8.2　关于一阶与二阶线性微分方程的结果的回顾

具有形式

$$y' + P(x)y = Q(x) \tag{8.1}$$

的微分方程叫作一阶线性微分方程, 其中 P 与 Q 为已知函数, y 为未知函数. 这是能用初等方法求解的少数几类微分方程中的一类. 下述定理给出了确定这类微分方程全部解的一个显式公式.

定理 8.1　设 P 与 Q 为开区间 J 上的连续函数, 则对 J 中任意一点 a 及任意实数 b, 存在唯一的函数 $y = f(x)$ 满足微分方程 (8.1) 和初始条件 $f(a) = b$, 并且 $f(x)$ 由下述显式公式给出,

$$f(x) = b\mathrm{e}^{-A(x)} + \mathrm{e}^{-A(x)} \int_a^x Q(t)\mathrm{e}^{A(t)}\,\mathrm{d}t, \tag{8.2}$$

其中 $A(x) = \int_a^x P(t)\,\mathrm{d}t$.

证明　在方程 (8.1) 两边同乘以 $\mathrm{e}^{A(x)}$, 由 $A'(x) = P(x)$ 可知, 方程 (8.1) 的左边变成 $y'\mathrm{e}^{A(x)} + y\mathrm{e}^{A(x)}A'(x)$, 这是乘积 $y\mathrm{e}^{A(x)}$ 的导数, 因此原方程可改写成

$$(y\mathrm{e}^{A(x)})' = Q(x)\mathrm{e}^{A(x)}.$$

我们把用来将方程 (8.1) 左边化为某个乘积函数的导数的因子 $\mathrm{e}^{A(x)}$ 叫作**积分因子**(integrating factor). 将上面这个方程在区间 $[a, x]$ 上求积, 则得

$$f(x)\mathrm{e}^{A(x)} - f(a)\mathrm{e}^{A(a)} = \int_a^x Q(t)\mathrm{e}^{A(t)}\,\mathrm{d}t.$$

由于 $f(a) = b$ 与 $A(a) = 0$, 代入上面等式即得方程 (8.1) 的解 (8.2). □

形如

$$y'' + P_1(x)y' + P_2(x)y = R(x)$$

的方程叫作**二阶线性微分方程**. 若系数 P_1 与 P_2 以及等式右边的 R 都在某个开区间 J 上连续, 则由一个存在性定理保证了这个方程在区间 J 上总有解存在. 然而并不存在类似于公式 (8.2) 那样用 P_1, P_2 和 R 来表示这些解的一般公式. 这样, 即使对方程 (8.1) 作这样相当简单的推广, 除了在某些特殊情形, 有关的理论远非完善. 不过当系数 P_1 与 P_2 为常数并且 $R = 0$ 时, 则下述定理指出, 此微分方程的解可以用多项式、指数函数和三角函数显式表示.

容易验证下述定理中所指出的函数都是这个微分方程的解, 而所有的解都包含在下面给出的公式 (8.4) 之中, 这一点则可作为定理 8.5 的推论.

定理 8.2　设 a 与 b 为给定的实常数, 令 $d = a^2 - 4b$, 则微分方程

$$y'' + ay' + by = 0 \tag{8.3}$$

在区间 $(-\infty, +\infty)$ 上的每一个解都可表成如下形式:

$$y = \mathrm{e}^{-ax/2}[c_1u_1(x) + c_2u_2(x)], \tag{8.4}$$

其中 c_1 与 c_2 为常数, 而函数 u_1 与 u_2 则依 d 的代数符号而不同:

(a) 若 $d = 0$, 则 $u_1(x) = 1$, $u_2(x) = x$.

(b) 若 $d > 0$, 则 $u_1(x) = \mathrm{e}^{kx}$, $u_2(x) = \mathrm{e}^{-kx}$, 此处 $k = \frac{1}{2}\sqrt{d}$.

(c) 若 $d < 0$, 则 $u_1(x) = \cos kx$, $u_2(x) = \sin kx$, 此处 $k = \frac{1}{2}\sqrt{-d}$. □

我们把 $d = a^2 - 4b$ 叫作二次方程

$$r^2 + ar + b = 0 \tag{8.5}$$

的**判别式**(discriminant). 通常将二次方程 (8.5) 叫作微分方程 (8.3) 的**特征方程**(characteristic equation), 它的根由公式

$$r_1 = \frac{-a+\sqrt{d}}{2} \quad 与 \quad r_2 = \frac{-a-\sqrt{d}}{2}$$

给出. d 的代数符号决定了这两个根的本性. 若 $d > 0$, 则 r_1 与 r_2 都是实数, 微分方程的解 (8.4) 由下式给出:

$$y = c_1\mathrm{e}^{r_1x} + c_2\mathrm{e}^{r_2x},$$

它是两个指数函数的线性组合. 若 $d < 0$, 则 r_1 与 r_2 是一对共轭复数:

$$r_1 = -\frac{1}{2}a + \mathrm{i}k, \quad r_2 = -\frac{1}{2}a - \mathrm{i}k, \quad 此处 k = \frac{1}{2}\sqrt{-d}.$$

在这一情形, 指数函数 e^{r_1x} 与 e^{r_2x} 是由下式给出的一对共轭的复函数:

$$\mathrm{e}^{r_1x} = \mathrm{e}^{-ax/2}\mathrm{e}^{\mathrm{i}kx} = \mathrm{e}^{-ax/2}\cos kx + \mathrm{i}\mathrm{e}^{-ax/2}\sin kx,$$

$$\mathrm{e}^{r_2x} = \mathrm{e}^{-ax/2}\mathrm{e}^{-\mathrm{i}kx} = \mathrm{e}^{-ax/2}\cos kx - \mathrm{i}\mathrm{e}^{-ax/2}\sin kx.$$

微分方程的解 (8.4) 是这两个复指数函数的实部和虚部的线性组合.

8.3 习　　题

本节习题用来复习关于一阶与二阶线性微分方程的基本内容.

一阶线性微分方程问题

在习题 1 ~ 3 中, 利用定理 8.1 求解在各区间 J 上的初值问题.

1. $y' - 3y = \mathrm{e}^{2x}$, $J = (-\infty, +\infty)$, 当 $x = 0$ 时 $y = 0$.
2. $xy' - 2y = x^5$, $J = (0, +\infty)$, 当 $x = 1$ 时 $y = 1$.
3. $y' + y\tan x = \sin 2x$, $J = (-\frac{1}{2}\pi, \frac{1}{2}\pi)$, 当 $x = 0$ 时 $y = 2$.
4. 某一类细菌以与其当前的总量成正比的速度繁殖并且 1 小时后细菌数加倍, 问它在 2 小时后增加了多少?
5. 一条过原点的曲线的笛卡儿方程为 $y = f(x)$. 过曲线上任意一点作两条坐标轴的平行线, 它们与坐标轴围成一个长方形. 曲线将这个长方形分隔成两个区域 A 与 B, 其中一个区域的面积是另一个区域面积的 n 倍, 求函数 f.
6. (a) 设 u 为二阶微分方程 $y'' + P(x)y' + Q(x)y = 0$ 的一个非零解.

   证明: 如果用 uv 代换微分方程

$$y'' + P(x)y' + Q(x)y = R(x)$$

中的 y, 那么该方程将变为关于 v' 的一个一阶线性方程.

(b) 由观察法求出方程 $y'' - 4y' + x^2(y' - 4y) = 0$ 的一个解, 然后用 (a) 中的方法求出方程

$$y'' - 4y' + x^2(y' - 4y) = 2xe^{-x^3/3}$$

的一个满足如下条件的解 y: 当 $x = 0$ 时 $y = 0$, $y' = 4$.

7. 作变量代换 $v = (y+1)^2$ 可将非线性微分方程

$$(y+1)y' + x(y^2 + 2y) = x$$

转换成一个关于 v 的一阶线性微分方程. 利用这一信息, 找出原非线性微分方程的所有解 $y = f(x)$.

常系数二阶线性微分方程

在习题 8~11 中, 找出微分方程在 $(-\infty, +\infty)$ 上的所有解:

8. $y'' - 4y = 0$.　　　　9. $y'' + 4y = 0$.

10. $y'' - 2y' + 5y = 0$.　　　　11. $y'' + 2y' + y = 0$.

12. (a) 求微分方程 $y'' - 2y' - 3y = 0$ 的满足初始条件 $f(0) = 5$ 和 $f'(0) = 35$ 的解 $y = f(x)$.
    (b) 求常数 A, B, C, 使得函数 $y = Ae^x + B\sin x + C\cos x$ 满足方程 $y'' - 2y' - 3y = 2e^x - 10\sin x$.

13. 求常数 k 的所有值, 使得微分方程 $y'' + ky = 0$ 有一个满足初始条件 $f_k(0) = f_k(1) = 0$ 的非平凡解 $y = f_k(x)$.

14. 设 (a, b) 是平面上给定的一点, m 是给定的一个实数. 证明: 微分方程 $y'' + k^2y = 0$ 恰有唯一一个图像通过点 (a, b) 并且在点 (a, b) 处斜率为 m 的解. $k = 0$ 的情形分开讨论.

15. 在下列各小题中, 求出一个二阶线性微分方程, 使得 u_1 与 u_2 为它的解.
    (a) $u_1(x) = e^x$,　$u_2(x) = e^{-x}$.
    (b) $u_1(x) = e^{2x}$,　$u_2(x) = xe^{2x}$.
    (c) $u_1(x) = e^{-x/2}\cos x$,　$u_2(x) = e^{-x/2}\sin x$.
    (d) $u_1(x) = \sin(2x+1)$,　$u_2(x) = \sin(2x+2)$.
    (e) $u_1(x) = \cosh x$,　$u_2(x)\sinh x$.
    注　$\cosh x = \frac{1}{2}(e^x + e^{-x})$, $\sinh x = \frac{1}{2}(e^x - e^{-x})$.

16. 一质点作简谐振动, 它在开始时的位移为 1, 速度为 1, 加速度为 -12, 求出质点在速度为 $\sqrt{8}$ 时的位移和加速度.

17. 给定微分方程 $y'' + P(x)y' + Q(x)y = 0$, 其中 P 与 Q 为处处连续的实函数.
    (a) 若方程有两个解 $y_1 = e^{ax}$ 与 $y_2 = e^{bx}$, 其中 a 与 b 为两个不同的实常数. 证明: P 与 Q 都是常数, 而且事实上我们有 $P(x) = -(a+b)$, $Q(x) = ab$.
    (b) 若设方程有两个解 $y_1 = e^{ax}$ 和 $y_2 = xe^{ax}$, 其中 a 为实常数, 证明或否定: P 与 Q 都是常数.

8.4　n 阶线性微分方程

若采用算子记号, 则可简化关于高阶线性微分方程的讨论. 令 $C(J)$ 表示区间 J 上全体实值连续函数组合的线性空间, 此处 J 可以是有界区间也可以是无界区间. 再设 f 是在 J 上的前 n 阶导数 $f', f'', \cdots, f^{(n)}$ 都存在并且连续的函数, 令 $C^{(n)}(J)$ 表示由所有这样的函数 f 组成的子空间, 令 $P_1, \cdots, P_n$ 为 $C(J)$ 中给定的 n 个函数, 考虑由

$$L(f) = f^{(n)} + P_1 f^{(n-1)} + \cdots + P_n f$$

所定义的算子 $L: C^{(n)}(J) \longrightarrow C(J)$, 有时算子 L 本身也表示为

$$L = D^n + P_1 D^{n-1} + \cdots + P_n,$$

其中 D^k 表示 k 阶微分算子. 采用算子记号后, n 阶线性微分方程可表示为

$$L(y) = R \tag{8.6}$$

的形式. 其中 R 是一个给定的定义在 J 上的函数. $C^{(n)}(J)$ 中任一在区间 J 上满足方程 (8.6) 的函数 y 叫作这个微分方程的一个解.

不难验证 $L(y_1 + y_2) = L(y_1) + L(y_2)$ 以及对任意常数 c 都有 $L(cy) = cL(y)$, 因此 L 叫作**线性算子**(linear operator), 这也是微分方程 $L(y) = R$ 叫作线性微分方程的原因. 这里的算子 L 叫作n **阶线性微分算子**.

对每个线性微分方程 $L(y) = R$, 将右边的函数 R 用零代替, 则得到如下微分方程

$$L(y) = 0,$$

上式叫作与 $L(y) = R$ 相对应的**齐次微分方程**. 当 R 不恒等于零时, 方程 $L(y) = R$ 叫作**非齐次微分方程**. 于是, 如果我们能解对应的齐次微分方程, 则我们总能解这个非齐次微分方程. 因此我们先处理齐次的情形.

齐次线性微分方程的解集是算子 L 的零空间 $N(L)$, 它也叫作这个方程的**解空间**(solution space). 解空间是 $C^{(n)}(J)$ 的子空间. 尽管 $C^{(n)}(J)$ 是无限维的, 但是解空间 $N(L)$ 总是有限维的, 事实上, 它的维数就等于算子 L 的阶 n, 这一结果是下述存在唯一性定理的一个推论.

8.5 存在唯一性定理

定理 8.3 (n 阶线性微分方程解的存在唯一性定理) *设 $P_1, \cdots, P_n$ 为开区间 J 上的 n 个连续函数, 令 L 为下述线性微分算子:*

$$L = D^n + P_1 D^{n-1} + \cdots + P_n.$$

若 $x_0 \in J$ 并且 $k_0, k_1, \cdots, k_{n-1}$ 为 n 个给定的实数, 则存在唯一的一个函数 $y = f(x)$, 它在 J 上满足齐次微分方程 $L(y) = 0$ 并且也满足初始条件

$$f(x_0) = k_0, \quad f'(x_0) = k_1, \quad \cdots, \quad f^{(n-1)}(x_0) = k_{n-1}. \qquad \square$$

注 $\mathbb{R}^n$ 中由 $(f(x_0), f'(x_0), \cdots, f^{(n-1)}(x_0))$ 给出的向量叫作 f 在 x_0 处的**初值向量**(initial-value vector). 定理 8.3 是说, 如果我们选择 J 中一点 x_0 和 $\mathbb{R}^n$ 中一个向量, 则齐次方程 $L(y) = 0$ 在 J 上存在唯一的一个解 $y = f(x)$, 它以这个向量作为在 x_0 点的初值向量. 特别地, 只有零解才会以零向量作为在任意点的初值向量. 当 $n = 2$ 时, 则恰有唯一的一个解, 它在给定的一点 x_0 处取给定的值 $f(x_0)$ 和给定的导数值 $f'(x_0)$. 例如, 若 $f(t)$ 表示作简谐振动的质点在时刻 t 时的位置, 则 $f(t)$ 由它在 t_0 时的初始位置和初始速度完全决定.

上述存在唯一性定理将作为第 10 章中关于微分方程组的更一般的存在唯一性定理的一个推论而得. 9.9 节中将给出关于常系数方程情形的一个证明.

8.6 齐次线性微分方程解空间的维数

定理 8.4 (维数定理) 设 $L: C^{(n)}(J) \longrightarrow C(J)$ 是由

$$L = D^n + P_1 D^{n-1} + \cdots + P_n \tag{8.7}$$

给出的一个 n 阶线性微分算子, 则方程 $L(y) = 0$ 的解空间的维数为 n.

证明 为说明线性代数中的有关定理是如何起作用的, 我们从存在唯一性定理来证明维数定理.

令 $\mathbb{R}^n$ 表示由实 n 序组构成的线性空间, 设 T 为将解空间 $N(L)$ 中的各个函数 f 映为 f 在 x_0 点的初值向量的线性变换, 即

$$T(f) = \big(f(x_0), f'(x_0), \cdots, f^{(n-1)}(x_0)\big),$$

此处 x_0 是 J 中取定的一点. 由唯一性定理知, 由 $T(f) = O$ 可得 $f = 0$. 因此, 由定理 4.10 可知 T 在 $N(L)$ 上是一一的, 从而 T^{-1} 也是从 $\mathbb{R}^n$ 到 $N(L)$ 上的一一映射. 因此由定理 4.11 即得 $\dim N(L) = \dim \mathbb{R}^n = n$. □

一旦知道了解空间的维数是 n, 就可取任意 n 个线性无关的解作为一组基, 从而作为维数定理的一个推论, 我们有下述定理.

定理 8.5 设 $L: C^{(n)}(J) \longrightarrow C(J)$ 为一个 n 阶线性微分算子, 若 $u_1, \cdots, u_n$ 为齐次线性微分方程 $L(y) = 0$ 在 J 上的 n 个线性无关的解, 则 J 上任意一个解 $y = f(x)$ 都可以表示成如下形式:

$$f(x) = \sum_{k=1}^{n} c_k u_k(x), \tag{8.8}$$

其中 $c_1, \cdots, c_n$ 为常数. □

注 由于微分方程 $L(y) = 0$ 的所有解都包含在公式 (8.8) 中, 因而此公式右边对于任意常数 $c_1, \cdots, c_n$ 的线性组合叫作微分方程的**通解**(general solution).

维数定理告诉我们, n 阶齐次线性微分方程的解空间总存在由 n 个解组成的一组基, 但它并没有告诉我们如何确定这组基. 事实上, 至今并未找到对每一个线性方程都能确定其解空间的一组基的简单方法. 不过, 对某些特殊类型的微分方程已经有一些特殊方法. 常系数微分方程就是这样的一类方程, 对这类微分方程, 确定其解空间的一组基的问题化为求一个 n 次多项式的根的问题.

8.7 常系数线性算子的代数

形如

$$A = a_0 \mathrm{D}^n + a_1 \mathrm{D}^{n-1} + \cdots + a_{n-1}\mathrm{D} + a_n \tag{8.9}$$

的线性算子叫作**常系数微分算子**或**常系数算子**(constant-coefficient operator), 其中 D 为求导算子, $a_0, a_1, \cdots, a_n$ 为实常数. 当 $a_0 \neq 0$ 时, 称算子 A 的**阶**为 n. 算子 A 可作用于在某个区间内各阶导数 $y', y'', \cdots, y^{(n)}$ 都存在的任意函数 y, 其结果是由

$$A(y) = a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y$$

给出的一个函数 $A(y)$.

在本节中, 我们将限于讨论在 $(-\infty, +\infty)$ 上任意阶导数都存在的函数, 所有这样的函数的集合记作 C^∞, 称为**无限次可微函数类**. 若 $y \in C^\infty$, 则显然也有 $A(y) \in C^\infty$.

关于线性变换的通常的代数运算 (加法、纯量乘法以及关于变换的复合的乘法) 同样适用于常系数算子. 两个 (不一定同阶的) 常系数算子 A 与 B 的和还是一个常系数算子, 同样, 纯量积 λA 也是常系数算子, 因此, 全体常系数算子的集合组成一个线性空间. A 与 B 的**积**(复合) 也仍是一个常系数算子, 因此常系数算子的和、积与纯量积满足为所有线性变换所满足的通常的交换律、结合律和分配律. 而且由于对所有正整数 r 与 s 都有 $\mathrm{D}^r\mathrm{D}^s = \mathrm{D}^s\mathrm{D}^r$, 因此任意两个常系数算子都可交换: $AB = BA$.

对于由等式 (8.9) 所给出的常系数算子 A, 我们令系数与 A 相同的一个多项式 P_A 与它相对应, 叫作 A 的**相伴多项式**(associated polynomial), 即

$$P_A(r) = a_0 r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n.$$

反过来, 给定一个实系数多项式 P, 存在一个与 P 具有相同系数的算子 A 与之对应. 下面定理指出, 算子与多项式之间的这个相伴关系是一个一一对应关系, 而且它把算子的和、积与纯量积对应到这些算子的相伴多项式的和、积与纯量积.

定理 8.6 设 A 与 B 表示两个分别以 P_A 与 P_B 为相伴多项式的常系数算子, 再令 λ 为任意实数, 则我们有下列结论.

(a) $A = B$, 当且仅当 $P_A = P_B$.

(b) $P_{A+B} = P_A + P_B$.

(c) $P_{AB} = P_A \cdot P_B$(多项式 P_A 与 P_B 的积).

(d) $P_{\lambda A} = \lambda P_A$

证明 先考虑 (a), 设 $P_A = P_B$, 由于 P_A 与 P_B 有相同的次数和相同的系数, 因此算子 A 与 B 有相同的阶和相同的系数, 即 $A = B$.

下面证明由 $A = B$ 可得 $P_A = P_B$. 等式 $A = B$ 表示 $A(y) = B(y)$ 对 C^∞ 中的所有 y 都成立. 取 $y = \mathrm{e}^{rx}$, 其中 r 为常数, 则得 $y^{(k)} = r^k \mathrm{e}^{rx}$ 对所有 $k \geqslant 0$ 都成立, 因此

$$A(y) = P_A(r)\mathrm{e}^{rx} \quad 及 \quad B(y) = P_B(r)\mathrm{e}^{rx}.$$

将它们代入等式 $A(y) = B(y)$ 并消去非零因子 e^{rx} 即得 $P_A(r) = P_B(r)$. 由于 r 是任意实数, 因此必有 $P_A = P_B$. 由此即得 (a).

由相伴多项式的定义立即得到 (b), (c) 与 (d). □

由定理 8.6 可知, 关于多项式 P_A 与 P_B 的与和、积以及纯量积相关的每一个代数关系对算子 A 与 B 同样成立. 特别地, 如果相伴多项式 P_A 可分解成两个或

更多个多项式的积, 那么每个因式必是某个常系数算子的相伴多项式, 因此, 由定理 8.6, 算子 A 也有一个相应的分解. 例如, 若 $P_A(r) = P_B(r)P_C(r)$, 则 $A = BC$. 若 $P_A(r)$ 可分解成 n 个线性因式的乘积, 设为

$$P_A(r) = a_0(r - r_1)(r - r_2)\cdots(r - r_n) \tag{8.10}$$

则 A 有相应的因式分解形式:

$$A = a_0(\mathrm{D} - r_1)(\mathrm{D} - r_2)\cdots(\mathrm{D} - r_n).$$

代数基本定理告诉我们, 每一个次数 $n \geqslant 1$ 的多项式 $P_A(r)$ 都有一个形如 (8.10) 的分解式, 其中 $r_1, \cdots, r_n$ 为它的根, 每一个根在这里出现的次数等于它的重数 (2 重根出现 2 次, 3 重根出现 3 次, 等等). 这些根可能是实数也可能是复数. 由于 $P_A(r)$ 的系数都是实数, 因此它的复根必以共轭复数的形式成对出现: $\alpha \pm \mathrm{i}\beta, \beta \neq 0$. 与每对共轭复根相对应的两个线性因式相乘给出一个实系数二次因式 $r^2 - 2\alpha r + \alpha^2 + \beta^2$. 于是, 每一个多项式 $P_A(r)$ 都可分解成实系数的线性因式与二次因式的乘积. 由此也给出了将算子 A 分解成实数系数的一阶与二阶常系数算子乘积的相应分解式.

例 1 设 $A = D^2-5D+6$, 相伴多项式 $P_A(r)$ 有分解式 $r^2-5r+6 = (r-2)(r-3)$, 因此算子 A 有分解式

$$\mathrm{D}^2 - 5\mathrm{D} + 6 = (\mathrm{D} - 2)(\mathrm{D} - 3).$$ □

例 2 设 $A = \mathrm{D}^4 - 2\mathrm{D}^3 + 2\mathrm{D}^2 - 2\mathrm{D} + 1$, 相伴多项式 $P_A(r)$ 有分解式

$$r^4 - 2r^3 + 2r^2 - 2r + 1 = (r - 1)(r - 1)(r^2 + 1),$$

因此 A 有分解式 $A = (\mathrm{D} - 1)(\mathrm{D} - 1)(\mathrm{D}^2 + 1)$. □

8.8 由算子的因式分解求常系数线性微分方程解的一组基

下述定理指出如何通过常系数算子的因式分解来求解常系数线性微分方程.

定理 8.7 *设 L 为常系数算子, L 可分解成如下常系数算子的积:*

$$L = A_1A_2\cdots A_k,$$

则线性微分方程 $L(y) = 0$ 的解空间包含了各个微分方程 $A_i(y) = 0$ 的解空间, 即

$$N(A_i) \subseteq N(L), \quad i = 1, 2, \cdots, k \tag{8.11}$$

证明 若 u 在最后一个因子 A_k 的零空间中, 则 $A_k(u) = O$, 因此 L 的零空间包含了最后一个因子 A_k 的零空间. 由于常系数算子是可交换的, 因此我们可以这样安排, 使得每一个因子都可以作为最后一个因子. 由此即得关系式 (8.11). □

若 $L(u) = O$, 则称算子 ***L*将 *u* 零化**(annihilate). 定理 8.7 告诉我们, 如果 L 的一个因子 A_i 将 u 零化, 则 L 也将 u 零化.

下列例子指出如何利用定理 8.7 解常系数齐次微分方程. 选择这些例子是用来说明微分方程的解依赖于相伴多项式根的性质的特征.

情形 1. 不同实根的情形

例 1 求微分方程

$$(\mathrm{D}^3 - 7\mathrm{D} + 6)y = 0 \tag{8.12}$$

的解的一组基.

解 这个方程具有 $L(y) = 0$ 的形式, 其中

$$L = \mathrm{D}^3 - 7\mathrm{D} + 6 = (\mathrm{D} - 1)(\mathrm{D} - 2)(\mathrm{D} + 3).$$

$(\mathrm{D} - 1)$ 的零空间包含解 $u_1(x) = \mathrm{e}^x$, $(\mathrm{D} - 2)$ 的零空间包含解 $u_2(x) = \mathrm{e}^{2x}$, $(\mathrm{D} + 3)$ 的零空间包含解 $u_3(x) = \mathrm{e}^{-3x}$. 在 3.7 节例 7 中我们已证明 u_1, u_2 与 u_3 这 3 个指数函数是线性无关的. 3 阶方程的任意 3 个线性无关的解组成解空间的一组基, 因此方程 (8.12) 的通解由

$$y = c_1\mathrm{e}^x + c_2\mathrm{e}^{2x} + c_3\mathrm{e}^{-3x}$$

给出. □

例 1 的解法可用来证明下述定理, 这个定理告诉我们, 对于任意一个能分解成不同线性因式乘积的常系数算子, 如何找出解空间的一组基.

定理 8.8 *设 L 为一个常系数算子, 它的相伴多项式 $P_L(r)$ 有 n 个不同的实根 $r_1, r_2, \cdots, r_n$, 则微分方程 $L(y) = 0$ 在区间 $(-\infty, +\infty)$ 上的通解由 n 个指数函数的线性组合给出:*

$$y = \sum_{k=1}^{n} c_k \mathrm{e}^{r_k x}. \tag{8.13}$$

证明 我们有如下因式分解:

$$L = a_0(\mathrm{D} - r_1)\cdots(\mathrm{D} - r_n).$$

由于 $(\mathrm{D} - r_k)$ 的零空间包含指数函数 $u_k(x) = \mathrm{e}^{r_k x}$, 因此 L 的零空间包含了下列 n 个指数函数:

$$u_1(x) = \mathrm{e}^{r_1 x}, \cdots, u_n(x) = \mathrm{e}^{r_n x}. \tag{8.14}$$

在 3.7 节例 7 中我们已证明这些函数是线性无关的, 因此它们组成微分方程 $L(y) = 0$ 的解空间的一组基, 于是通解由这些函数的线性组合 (8.13) 给出. □

情形 2. 都是实根, 但有重根

如果所有的根都是实数但有重根, 则 (8.14) 中的函数不是线性无关的, 因此它们不组成解空间的一组基. 如果某个根 r 出现 m 次, 则 $(\mathrm{D} - r)^m$ 是 L 的一个因式. 下面的定理告诉我们如何从这个因式的零空间中求得 m 个线性无关的解.

定理 8.9 *m 个函数*

$$u_1(x) = \mathrm{e}^{rx}, \quad u_2(x) = x\mathrm{e}^{rx}, \quad \cdots, \quad u_m(x) = x^{m-1}\mathrm{e}^{rx}$$

线性无关并且都被算子 $(\mathrm{D} - r)^m$ 零化.

证明 这些函数的线性无关性由多项式 $1, x, x^2, \cdots, x^{m-1}$ 的线性无关性导出. 为证明 $u_1, \cdots, u_m$ 被 $(\mathrm{D} - r)^m$ 零化, 我们对 m 作归纳法.

若 $m = 1$, 则仅有 $u_1(x) = \mathrm{e}^{rx}$ 一个函数, 它当然被 $(\mathrm{D} - r)$ 零化. 假设定理对 $m - 1$ 成立, 即函数 $u_1, \cdots, u_{m-1}$ 都被 $(\mathrm{D} - r)^{m-1}$ 零化, 由于

$$(\mathrm{D}-r)^m=(\mathrm{D}-r)(\mathrm{D}-r)^{m-1},$$

因此, 函数 $u_1,\cdots,u_{m-1}$ 也都被 $(\mathrm{D}-r)^m$ 零化. 为完成定理的证明, 我们必须证明 $(\mathrm{D}-r)^m$ 也将 u_m 零化. 考察

$$(\mathrm{D}-r)^m u_m=(\mathrm{D}-r)^{m-1}(\mathrm{D}-r)(x^{m-1}\mathrm{e}^{rx}).$$

我们有

$$\begin{aligned}(\mathrm{D}-r)(x^{m-1}\mathrm{e}^{rx})&=\mathrm{D}(x^{m-1}\mathrm{e}^{rx})-rx^{m-1}\mathrm{e}^{rx}\\&=(m-1)x^{m-2}\mathrm{e}^{rx}+x^{m-1}r\mathrm{e}^{rx}-rx^{m-1}\mathrm{e}^{rx}\\&=(m-1)x^{m-2}\mathrm{e}^{rx}=(m-1)u_{m-1}(x).\end{aligned}$$

将 $(\mathrm{D}-r)^{m-1}$ 作用于上式两端, 由于 $(\mathrm{D}-r)^{m-1}$ 将 u_{m-1} 零化, 因此右端化为 0, 从而 $(\mathrm{D}-r)^m u_m=0$, 即 u_m 被 $(\mathrm{D}-r)^m$ 零化. 即得定理. □

例 2　求微分方程 $L(y)=0$ 的通解, 此处

$$L=\mathrm{D}^3-\mathrm{D}^2-8\mathrm{D}+12.$$

解　算子 L 有因式分解

$$L=(\mathrm{D}-2)^2(\mathrm{D}+3).$$

由定理 8.9, 函数

$$u_1(x)=\mathrm{e}^{2x},\quad u_2(x)=x\mathrm{e}^{2x}$$

都在 $(\mathrm{D}-2)^2$ 的零空间内. 函数 $u_3(x)=\mathrm{e}^{-3x}$ 在 $(\mathrm{D}+3)$ 的零空间内. 由于 u_1,u_2,u_3 线性无关 (见 8.9 节习题 17), 因此他们组成 L 的零空间的一组基, 于是微分方程的通解为

$$y=c_1\mathrm{e}^{2x}+c_2x\mathrm{e}^{2x}+c_3\mathrm{e}^{-3x}.$$ □

定理 8.9 告诉我们, 当一个 n 阶常系数线性微分方程的相伴多项式的根全是实数但有重根时如何找出这个方程的解空间的一组基. 如果不同的根为 $r_1,\cdots,r_k$, 它们的重数分别为 $m_1,\cdots,m_k$, 则在解空间的一组基中与根 r_p 相对应的 m_p 个函数为

$$u_{q,p}(x)=x^{q-1}\mathrm{e}^{r_px},\quad \text{此处}\, q=1,2,\cdots,m_p.$$

当 p 取遍 $1,2,\cdots,k$ 时, 我们共得到 $m_1+\cdots+m_k$ 个函数. 在 8.9 节习题 17 中, 我们还简略地证明了所有这些函数是线性无关的. 由于 $m_1+\cdots+m_k=n$(即微分方程的阶), 因此 n 个函数 $u_{p,q}$ 组成此微分方程解空间的一组基.

例 3　解微分方程 $(\mathrm{D}^6+2\mathrm{D}^5-2\mathrm{D}^3-\mathrm{D}^2)y=0$.

解　我们有 $\mathrm{D}^6+2\mathrm{D}^5-2\mathrm{D}^3-\mathrm{D}^2=\mathrm{D}^2(\mathrm{D}-1)(\mathrm{D}+1)^3$. 解空间的基中与因式 D^2 相对应的函数为 $u_1(x)=1$ 与 $u_2(x)=x$, 与因式 $(\mathrm{D}-1)$ 相对应的函数为 $u_3(x)=\mathrm{e}^x$; 与因式 $(\mathrm{D}+1)^3$ 相对应的函数为 $u_4(x)=\mathrm{e}^{-x}$, $u_5(x)=x\mathrm{e}^{-x}$ 及 $u_6(x)=x^2\mathrm{e}^{-x}$. $u_1,\cdots,u_6$ 这 6 个函数线性无关, 因此方程的通解为

$$y=c_1+c_2x+c_3\mathrm{e}^x+(c_4+c_5x+c_6x^2)\mathrm{e}^{-x}.$$ □

情形 3. 有复数根

如用到复指数函数, 则无需区分算子 L 的相伴多项式的根为实数还是复数. 如果只需求实值的解, 那么我们将 L 分解成实系数的一次与二次因式. 任意一对共轭复根 $\alpha \pm i\beta$ 对应于一个二次因式

$$D^2 - 2\alpha D + \alpha^2 + \beta^2. \tag{8.15}$$

这个二阶算子的零空间包含两个线性无关的函数 $u(x) = e^{\alpha x}\cos\beta x$ 与 $v(x) = e^{\alpha x}\sin\beta x$. 如果共轭复根 $\alpha \pm i\beta$ 的重数为 m, 则此二次因式的幂次为 m, 算子

$$(D^2 - 2\alpha D + \alpha^2 + \beta^2)^m$$

的零空间包含如下 $2m$ 个线性无关的函数:

$$u_q(x) = x^{q-1}e^{\alpha x}\cos\beta x, \quad v_q(x) = x^{q-1}e^{\alpha x}\sin\beta x, \quad q = 1, 2, \cdots, m.$$

这可通过对 m 作归纳法得到证明. 下面的一个例子说明了某些可能的情形.

例 4　$y''' - 4y'' + 13y' = 0$, 相伴多项式 $r^3 - 4r^2 + 13r$ 的根为 $0, 2 \pm 3i$, 通解为

$$y = c_1 + e^{2x}(c_2\cos 3x + c_3\sin 3x).$$ □

例 5　$y''' - 2y'' + 4y' - 8y = 0$, 相伴多项式为

$$r^3 - 2r^2 + 4r - 8 = (r-2)(r^2+4),$$

它的根是 2, $\pm 2i$, 因此微分方程的通解为

$$y = c_1e^{2x} + c_2\cos 2x + c_3\sin 2x.$$ □

例 6　$y^{(5)} - 9y^{(4)} + 34y''' - 66y'' + 65y' - 25y = 0$, 相伴多项式可分解成

$$(r-1)(r^2-4r+5)^2,$$

它的根为 1, $2 \pm i$, $2 \pm i$. 因此微分方程的通解是

$$y = c_1e^x + e^{2x}\{(c_2 + c_3x)\cos x + (c_4 + c_5x)\sin x\}.$$ □

8.9 习　题

在习题 1 ~ 12 中, 求各微分方程的通解.

1. $y''' - 2y'' - 3y' = 0$.
2. $y''' - y' = 0$.
3. $y''' + 4y'' + 4y' = 0$.
4. $y''' - 3y'' + 3y' - y = 0$.
5. $y^{(4)} + 4y''' + 6y'' + 4y' + y = 0$.
6. $y^{(4)} - 16y = 0$.
7. $y^{(4)} + 16y = 0$.
8. $y''' - y = 0$.
9. $y^{(4)} + 4y''' + 8y'' + 8y' + 4y = 0$.
10. $y^{(4)} + 2y'' + y = 0$.
11. $y^{(6)} + 4y^{(4)} + 4y'' = 0$.
12. $y^{(6)} + 8y^{(4)} + 16y'' = 0$.
13. 设 m 为正常数, 求微分方程

$$y''' - my'' + m^2y' - m^3y = 0$$

的满足条件 $f(0) = f'(0) = 0$, $f''(0) = 1$ 的特解 $y = f(x)$.

14. 设某个常系数线性微分方程的相伴多项式为 $f(r)$, 若 $f(r)$ 的所有根都是负实数, 证明: 微分方程的每一个解当 $x \longrightarrow +\infty$ 都趋于 0. 若相伴多项式的所有根都是非正实数, 则对方程在区间 $(0, +\infty)$ 上的所有解的性状你能得出什么结论?

15. 在下列各小题中, 求出给定函数所满足的一个常系数线性微分方程.

(a) $u_1(x)=e^x$,　$u_2(x)=e^{-x}$,　$u_3(x)=e^{2x}$,　$u_4(x)=e^{-2x}$.

(b) $u_1(x)=e^{-2x}$,　$u_2(x)=xe^{-2x}$,　$u_3(x)=x^2e^{-2x}$.

(c) $u_1(x)=1$,　$u_2(x)=x$,　$u_3(x)=e^x$,　$u_4(x)=xe^x$.

(d) $u_1(x)=x$,　$u_2(x)=e^x$,　$u_3(x)=xe^x$.

(e) $u_1(x)=x^2$,　$u_2(x)=e^x$,　$u_3(x)=xe^x$.

(f) $u_1(x)=e^{-2x}\cos 3x$,　$u_2(x)=e^{-2x}\sin 3x$,　$u_3(x)=e^{-2x}$,　$u_4(x)=xe^{-2x}$.

(g) $u_1(x)=\cosh x$,　$u_2(x)=\sinh x$,　$u_3(x)=x\cosh x$,　$u_4(x)=x\sinh x$.

(h) $u_1(x)=\cosh x\sin x$,　$u_2(x)=\sinh x\cos x$,　$u_3(x)=x$.

16. 设 $r_1,\cdots,r_n$ 为 n 个不同实数, $Q_1,\cdots,Q_n$ 为 n 个非零多项式. 证明: n 个函数

$$u_1(x)=Q_1(x)e^{r_1x},\quad\cdots,\quad u_n(x)=Q_n(x)e^{r_nx}$$

线性无关.

证明要点　对 n 作归纳法. 当 $n=1$ 与 $n=2$ 时容易验证结论. 假定当 $n=p$ 时结论成立, 令 $c_1,\cdots,c_{p+1}$ 为 $p+1$ 个纯量, 使得

$$\sum_{k=1}^{p+1}c_kQ_k(x)e^{r_kx}=0.$$

将上式两边同乘以 $e^{-r_{p+1}x}$ 之后再求导, 然后利用归纳假设证明所有的系数 c_k 都等于零. 另一个证明可以像 3.7 节例 7 那样由 $x\to+\infty$ 的数量级给出.

17. 设 $m_1,m_2,\cdots,m_k$ 为 k 个正整数, $r_1,r_2,\cdots,r_k$ 为 k 个不同的实数, 再令 $n=m_1+\cdots+m_k$. 对每一对满足条件 $1\leqslant p\leqslant k$, $1\leqslant q\leqslant m_p$ 的整数 p,q, 令

$$u_{q,p}(x)=x^{q-1}e^{r_px}.$$

例如, 当 $p=1$ 时, 相应的函数为

$$u_{1,1}(x)=e^{r_1x},\quad u_{2,1}(x)=xe^{r_1x},\quad\cdots,\quad u_{m_1,1}(x)=x^{m_1-1}e^{r_1x}.$$

证明: 这样定义的 n 个函数 $u_{q,p}(x)$ 线性无关. [提示: 利用习题 16.]

8.10　齐次方程与非齐次方程之间的关系

我们回到系数不一定为常数的一般 n 阶线性微分方程上来. 下述定理刻划了齐次方程 $L(y)=0$ 的解与非齐次方程 $L(y)=R(x)$ 的解之间的关系.

定理 8.10　*设 $L:C^{(n)}(J)\longrightarrow C(J)$ 为一个 n 阶线性微分算子. 设 $u_1,\cdots,u_n$ 为齐次方程 $L(y)=0$ 的 n 个线性无关的解, 再设 y_1 为非齐次方程 $L(y)=R$ 的一个特解, 此处 $R\in C(J)$, 则非齐次方程 $y=f(x)$ 的每一个解都可表成*

$$f(x)=y_1(x)+\sum_{k=1}^{n}c_ku_k(x)\tag{8.16}$$

的形式, 其中 $c_1,c_2,\cdots,c_n$ 为常数.

证明　设 f 为 $L(y)=R$ 的任意一个解. 由线性性得 $L(f-y_1)=L(f)-L(y_1)=R-R=0$. 因此 $f-y_1$ 是齐次方程 $L(y)=0$ 的解空间中的一个函数, 从而 $f-y_1$ 是 $u_1,\cdots,u_n$ 的一个线性组合, 即得等式 (8.16).　□

由于 $L(y)=R$ 的所有解都可从 (8.16) 中找到, 因此等式 (8.16) 右边的和式 (对于任意常数 $c_1,\cdots,c_n$) 叫作非齐次方程 $L(y)=R$ 的通解. 定理 8.10 告诉我们, 非齐次方程的通解可以由它的一个特解 y_1 加上对应的齐次方程的通解得到.

注 定理 8.10 有一个简单的几何类比, 它可以帮助我们理解定理的意义. 为确定一个平面上的所有点, 我们只需找出此平面上的一个特定的点, 然后将过原点且与原平面平行的平面上的点与这个点相加即可. 为找到 $L(y)=R$ 的全部解, 我们先找出它的一个特解, 然后将它与齐次方程 $L(y)=0$ 的所有解相加即可. 非齐次方程的解的集合相当于过一特定点的平面, 对应齐次方程的解空间类似于过原点的一个平行平面.

在实际应用定理 8.10 时, 我们必须解两个问题: (1) 求出齐次方程 $L(y)=0$ 的通解; (2) 求出非齐次方程 $L(y)=R$ 的一个特解. 下一节将证明, 只要我们能解问题 (1), 就也能解问题 (2).

8.11 求非齐次方程的一个特解 · 参数变易法

现在我们转到寻求非齐次方程 $L(y)=R$ 的一个特解 y_1 的问题. 我们将详细介绍一个叫作**参数变易法**(variation of parameters) 的十分有力的方法, 这个方法告诉我们在知道齐次方程 $L(y)=0$ 的 n 个线性无关的特解 $u_1,\cdots,u_n$ 时如何确定 y_1. 这个方法给出了形如

$$y_1=v_1u_1+\cdots+v_nu_n \tag{8.17}$$

的一个特解, 其中 $v_1,\cdots,v_n$ 为 n 个函数, 它们可以从 $u_1,\cdots,u_n$ 和方程 $L(y)=R$ 右边的函数 R 求出. 用这个方法将得到导数 $v_1',\cdots,v_n'$ 所满足的 n 个线性代数方程所组成的方程组. 由于这个线性方程组的系数矩阵为非奇异矩阵, 因此它总是有解的. 再将这些导数积分就得到所求的函数 $v_1,\cdots,v_n$. 参数变易法最早被 Johann Bernoulli 用来解一阶线性微分方程, 后来 Lagrange 于 1774 年用这个方法解二阶线性微分方程.

对于 n 阶的情形, 若采用向量和矩阵记号, 则可使表述得以简化. (8.17) 的右边可写成内积的形式:

$$y_1=(\boldsymbol{v},\boldsymbol{u}), \tag{8.18}$$

此处 v 与 u 分别是由

$$\boldsymbol{v}=(v_1,\cdots,v_n) \quad 与 \quad \boldsymbol{u}=(u_1,\cdots,u_n)$$

给出的 n 维向量函数.

将一个向量值函数的各个分量求导后组成的向量叫作它的导数, 例如, $\boldsymbol{v}'=(v_1',\cdots,v_n')$ 及 $\boldsymbol{u}'=(u_1',\cdots,u_n')$.

关于实值函数的和的导数及纯量倍的导数的熟知法则对向量值函数的和与纯量倍也同样适用. 乘法法则对内积也适用, 即 $(\boldsymbol{v},\boldsymbol{u})'$ 等于两个内积的和: $(\boldsymbol{v},\boldsymbol{u})'=(\boldsymbol{v},\boldsymbol{u}')+(\boldsymbol{v}',\boldsymbol{u})$.

在参数变易法中, 我们要选取 v 使得当 $L(\boldsymbol{u})=\boldsymbol{O}$ 时, 等式 (8.18) 中的内积 $y_1=(\boldsymbol{v},\boldsymbol{u})$ 满足非齐次方程 $L(y)=R$, 此处 $L(\boldsymbol{u})$ 是以 $L(u_1),\cdots,L(u_n)$ 为分量的

向量. 我们从计算 y_1 的一阶导数开始, 由乘法法则得

$$y_1' = (\boldsymbol{v}, \boldsymbol{u}') + (\boldsymbol{v}', \boldsymbol{u}). \tag{8.19}$$

总共有 n 个待定函数 $v_1, \cdots, v_n$, 因此我们可对这些函数提出 n 个条件. 如果要求等式 (8.19) 右边第二项为零, 则关于 y_1' 的公式可简化为

$$y_1' = (\boldsymbol{v}, \boldsymbol{u}'), \quad 若\,(\boldsymbol{v}', \boldsymbol{u}) = 0.$$

将上式两边微分, 得

$$y_1'' = (\boldsymbol{v}, \boldsymbol{u}'') + (\boldsymbol{v}', \boldsymbol{u}').$$

如果我们能选取 $\boldsymbol{v}$ 使得 $(\boldsymbol{v}', \boldsymbol{u}') = 0$, 则关于 y_1'' 的公式也可简化为如下形式:

$$y_1'' = (\boldsymbol{v}, \boldsymbol{u}''), \quad 若\,(\boldsymbol{v}', \boldsymbol{u}') = 0.$$

继续这一步骤, 直到 y_1 的 $n-1$ 阶导数, 我们得

$$y_1^{(n-1)} = (\boldsymbol{v}, \boldsymbol{u}^{(n-1)}), \quad 若\,(\boldsymbol{v}', \boldsymbol{u}^{(n-2)}) = 0.$$

到目前为止, 我们只给出了关于 $\boldsymbol{v}$ 的 $n-1$ 个条件, 再求一次导数, 由乘法法则得

$$y_1^{(n)} = (\boldsymbol{v}, \boldsymbol{u}^{(n)}) + (\boldsymbol{v}', \boldsymbol{u}^{(n-1)}).$$

这一次我们提出的条件是 $(\boldsymbol{v}', \boldsymbol{u}^{(n-1)}) = R(x)$, 于是上式变为

$$y_1^{(n)} = (\boldsymbol{v}, \boldsymbol{u}^{(n)}) + R(x), \quad 若\,(\boldsymbol{v}', \boldsymbol{u}^{(n-1)}) = R(x).$$

现在假定加到 $\boldsymbol{v}$ 上的 n 个条件都满足, 设 $L = \mathrm{D}^n + P_1(x)\mathrm{D}^{n-1} + \cdots + P_n(x)$, 将 L 作用于 y_1 上, 则得

$$\begin{aligned} L(y_1) &= y_1^{(n)} + P_1(x)y_1^{(n-1)} + \cdots + P_n(x)y_1 \\ &= \left\{(\boldsymbol{v}, \boldsymbol{u}^{(n)}) + R(x)\right\} + P_1(x)(\boldsymbol{v}, \boldsymbol{u}^{(n-1)}) + \cdots + P_n(x)(\boldsymbol{v}, \boldsymbol{u}) \\ &= (\boldsymbol{v}, L(\boldsymbol{u})) + R(x) = (\boldsymbol{v}, \boldsymbol{O}) + R(x) = R(x). \end{aligned}$$

因此 $L(y_1) = R(x)$. 于是 y_1 是非齐次方程的一个解.

如果关于 $\boldsymbol{v}$ 的 n 个条件都能满足, 则上述方法有效. 这些条件是 $(\boldsymbol{v}', \boldsymbol{u}^{(k)}) = 0$ 对 $k = 0, 1, \cdots, n-2$ 都成立以及 $(\boldsymbol{v}', \boldsymbol{u}^{(n-1)}) = R(x)$. 我们可将这 n 个条件用单独一个矩阵方程表示:

$$\boldsymbol{W}(x)\boldsymbol{v}'(x) = R(x)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{8.20}$$

此处 $\boldsymbol{v}'(x)$ 可看作一个 $n \times 1$ 列矩阵, $\boldsymbol{W}$ 是以 $\boldsymbol{u}$ 及其各阶导数的分量作为行向量的 $n \times n$ 矩阵:

$$\boldsymbol{W} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ u_1' & u_2' & \cdots & u_n' \\ \vdots & \vdots & \vdots & \vdots \\ u_1^{(n-1)} & u_2^{(n-1)} & \cdots & u_n^{(n-1)} \end{bmatrix}.$$

为纪念数学家 J. M. H. Wronski(1778 − 1853), 上述矩阵叫作 $u_1, \cdots, u_n$ 的**Wronski 矩阵**.

在下一节中我们将证明, 在 $u_1, \cdots, u_n$ 线性无关的情形, 其 Wronski 矩阵是非奇异矩阵. 因此, 在等式 (8.20) 两边同乘以 $\boldsymbol{W}(x)$ 的逆矩阵, 得

$$\boldsymbol{v}'(x) = R(x)\boldsymbol{W}(x)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}.$$

在所考察的区间 J 内选择点 c 与点 x 并将上述向量方程在从 c 到 x 的区间上 (依分量逐个) 积分, 则得

$$\boldsymbol{v}(x) - \boldsymbol{v}(c) = \int_c^x R(t)\boldsymbol{W}(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}\mathrm{d}t = \boldsymbol{z}(x),$$

此处

$$\boldsymbol{z}(x) = \int_c^x R(t)\boldsymbol{W}(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}\mathrm{d}t.$$

于是 $\boldsymbol{v}(x) = \boldsymbol{v}(c) + \boldsymbol{z}(x)$, 关于特解的表示式 $y_1 = (\boldsymbol{u}, \boldsymbol{v})$ 变为

$$y_1 = (\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}(c) + \boldsymbol{z}) = (\boldsymbol{u}, \boldsymbol{v}(c)) + (\boldsymbol{u}, \boldsymbol{z}).$$

上式右边第一项 $(\boldsymbol{u}, \boldsymbol{v}(c))$ 是 $u_1, \cdots, u_n$ 的一个线性组合, 因此它是齐次方程的解. 于是我们可以舍去这一项而用第二项 $(\boldsymbol{u}, \boldsymbol{z})$ 作为非齐次方程的一个特解. 换言之, $L(y) = R(x)$ 的一个特解可以由内积

$$(\boldsymbol{u}(x), \boldsymbol{z}(x)) = \left(\boldsymbol{u}(x), \int_c^x R(t)\boldsymbol{W}(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}\mathrm{d}t\right)$$

给出, 我们将上述讨论总结成下述定理.

定理 8.11 令 $u_1, \cdots, u_n$ 为区间 J 上 n 阶齐次线性微分方程 $L(y) = 0$ 的 n 个线性无关的特解, 则非齐次方程 $L(y) = R$ 的一个特解 y_1 由公式

$$y_1 = v_1 u_1 + \cdots + v_n u_n$$

给出, 此处 $v_1, \cdots, v_n$ 是由

$$\boldsymbol{v}(x) = \int_c^x R(t)\boldsymbol{W}(t)^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}\mathrm{d}t \tag{8.21}$$

给出的 $n\times 1$ 列矩阵 v 的各个分量. 在上述表示式中, $\boldsymbol{W}$ 是 $u_1,\cdots,u_n$ 的 Wronski 矩阵, c 为 J 中任意一点. □

例 1　求微分方程

$$y''-y=\frac{2}{1+\mathrm{e}^x}$$

在区间 $(-\infty,+\infty)$ 上的通解.

解　齐次方程 $(\mathrm{D}^2-1)y=0$ 有两个线性无关的解 $u_1(x)=\mathrm{e}^x$ 与 $u_2(x)=\mathrm{e}^{-x}$. u_1 与 u_2 的 Wronski 矩阵为

$$\boldsymbol{W}(x)=\begin{bmatrix}\mathrm{e}^x & \mathrm{e}^{-x}\\ \mathrm{e}^x & -\mathrm{e}^{-x}\end{bmatrix},$$

其行列式等于 -2. 因此 $\boldsymbol{W}(x)$ 非奇异且其逆为

$$\boldsymbol{W}(x)^{-1}=-\frac{1}{2}\begin{bmatrix}-\mathrm{e}^{-x} & -\mathrm{e}^{-x}\\ -\mathrm{e}^x & \mathrm{e}^x\end{bmatrix}.$$

于是

$$\boldsymbol{W}(x)^{-1}\begin{bmatrix}0\\1\end{bmatrix}=-\frac{1}{2}\begin{bmatrix}-\mathrm{e}^{-x}\\ \mathrm{e}^x\end{bmatrix},$$

因此我们有

$$R(t)\boldsymbol{W}(t)^{-1}\begin{bmatrix}0\\1\end{bmatrix}=-\frac{1}{2}\frac{2}{1+\mathrm{e}^t}\begin{bmatrix}-\mathrm{e}^{-t}\\ \mathrm{e}^t\end{bmatrix}=\begin{bmatrix}\dfrac{\mathrm{e}^{-t}}{1+\mathrm{e}^t}\\ \dfrac{-\mathrm{e}^t}{1+\mathrm{e}^t}\end{bmatrix}.$$

将上式右边向量的各分量从 0 到 x 积分得

$$\begin{aligned}v_1(x)&=\int_0^x\frac{\mathrm{e}^{-t}}{1+\mathrm{e}^t}\mathrm{d}t=\int_0^x\left(\mathrm{e}^{-t}-1+\frac{\mathrm{e}^t}{1+\mathrm{e}^t}\right)\mathrm{d}t\\&=-\mathrm{e}^{-x}-x+\ln(1+\mathrm{e}^x)+1-\ln 2,\\ v_2(x)&=\int_0^x\frac{-\mathrm{e}^t}{1+\mathrm{e}^t}\mathrm{d}t=-\ln(1+\mathrm{e}^x)+\ln 2.\end{aligned}$$

由于关于 $v_1(x)$ 与 $v_2(x)$ 的上述表示式中的常数项只与齐次方程的解有关, 因此可以舍去, 于是我们得到非齐次方程的一个特解

$$y_1=v_1u_1+v_2u_2=-1-x\mathrm{e}^x+(\mathrm{e}^x-\mathrm{e}^{-x})\ln(1+\mathrm{e}^x).$$

从而非齐次方程的通解为

$$y=y_1+c_1\mathrm{e}^x+c_2\mathrm{e}^{-x}.$$

□

8.12　齐次线性微分方程 n 个线性无关解的 Wronski 矩阵的非奇异性

在本节中, 我们要证明 n 阶齐次方程 $L(y)=0$ 的 n 个线性无关解 $u_1,\cdots,u_n$ 的 Wronski 矩阵 $\boldsymbol{W}$ 是非奇异的. 为此我们将证明它的行列式是一个指数函数, 因

而在所论及的区间 J 上其值不可能为零, 由此即得此 Wronski 矩阵非奇异.

对任意 $x \in J$, 令 $w(x) = \det \boldsymbol{W}(x)$, 假定 $u_1, \cdots, u_n$ 所满足的微分方程为下述形式:

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = 0, \tag{8.22}$$

则我们有下述结论.

定理 8.12 Wronski 行列式 $w(x)$ 在 J 上满足下述一阶微分方程:

$$w' + P_1(x)w = 0, \tag{8.23}$$

从而当 $c \in J$ 时, 我们有

$$w(x) = w(c)\exp\left(-\int_c^x P_1(t)\,\mathrm{d}t\right) \quad (\text{Abel 公式}). \tag{8.24}$$

而且对所有 J 中的点 x, 都有 $w(x) \neq 0$.

证明 令 $\boldsymbol{u}$ 为行向量 $\boldsymbol{u} = (u_1, \cdots, u_n)$. 由于 $\boldsymbol{u}$ 的各个分量都满足微分方程 (8.22), 所以 $\boldsymbol{u}$ 也满足方程 (8.22). Wronski 矩阵 $\boldsymbol{W}$ 的各行是向量 $\boldsymbol{u}, \boldsymbol{u}', \cdots, \boldsymbol{u}^{(n-1)}$, 因此其行列式 w 可表成

$$w = \det \boldsymbol{W} = \det(\boldsymbol{u}, \boldsymbol{u}', \cdots, \boldsymbol{u}^{(n-2)}, \boldsymbol{u}^{(n-1)}).$$

w 的导数是将 $\boldsymbol{W}$ 最后一行求导所得矩阵的行列式 (见 5.22 节习题 7), 即

$$w' = \det(\boldsymbol{u}, \boldsymbol{u}', \cdots, \boldsymbol{u}^{(n-2)}, \boldsymbol{u}^{(n)}).$$

将 w 中最后一行乘以 $P_1(x)$, 我们得

$$P_1(x)w = \det(\boldsymbol{u}, \boldsymbol{u}', \cdots, \boldsymbol{u}^{(n-2)}, P_1(x)\boldsymbol{u}^{(n-1)}).$$

将上述两式相加得

$$w' + P_1(x)w = \det(\boldsymbol{u}, \boldsymbol{u}', \cdots, \boldsymbol{u}^{(n-2)}, \boldsymbol{u}^{(n)} + P_1(x)\boldsymbol{u}^{(n-1)}).$$

由于 u 满足微分方程 (8.22), 因此上式中的行列式的行向量组线性相关, 因此行列式等于零, 这意味着 w 满足方程 (8.23). 求解一阶微分方程 (8.23) 即得 Abel 公式 (8.24).

为证明 $w(x)$ 在 J 上不为零, 只需证明对 J 中某点 c 有 $w(c) \neq 0$ 即可. 我们用反证法证明这一结论. 假定对 J 中所有的 t 都有 $w(t) = 0$, 在 J 中选定一点 t, 设 $t = t_0$, 考虑齐次线性方程组

$$\boldsymbol{W}(t_0)\boldsymbol{X} = \boldsymbol{O},$$

此处 $\boldsymbol{X}$ 是一个列向量. 由于 $w(t_0) = \det W(t_0) = 0$, 因此 $\boldsymbol{W}(t_0)$ 为奇异矩阵, 从而上述齐次线性方程组有一个非零解向量, 设 $\boldsymbol{X}^{\mathrm{T}} = (c_1, \cdots, c_n) \neq (0, \cdots, 0)$. 以此向量的分量为系数作线性组合

$$f(t) = c_1u_1(t) + \cdots + c_nu_n(t).$$

如此定义的函数 f 在 J 上满足齐次方程 $L(f) = 0$, 这是因为 f 是 $u_1, \cdots, u_n$ 的线性组合. 由矩阵方程 $\boldsymbol{W}(t_0)\boldsymbol{X} = \boldsymbol{O}$ 得

$$f(t_0) = f'(t_0) = \cdots = f^{(n-1)}(t_0) = 0.$$

因此 f 在 $t=t_0$ 处的初值向量为零向量 $\boldsymbol{O}$, 由唯一性定理, f 为零解, 即 $(c_1,\cdots,c_n)=(0,\cdots,0)$, 与假设 $(c_1,\cdots,c_n)$ 为非零向量矛盾. 因此对 J 中某个 t 有 $w(t)\neq 0$. 在 Abel 公式中取 c 为这个 t, 则可见 $w(x)$ 在 J 上不可能为零. 由此即得定理 8.12. □

8.13　求非齐次方程特解的特殊方法・化为一阶线性微分方程组

由于参数变易法是求非齐次方程特解的一个普遍适用的方法, 因此在某些特殊情形下它不一定是最简便的方法. 当方程具有某种特殊形式时, 常常会有某些较为便捷的特殊解法. 例如当方程的系数为常数时, 我们可以把问题化为一系列一阶线性微分方程的问题, 这个方法最适宜于用一个简单例子说明.

例　求二阶微分方程

$$(\mathrm{D}-1)(\mathrm{D}-2)y=x\mathrm{e}^{x+x^2} \tag{8.25}$$

的一个特解.

解　令 $u=(\mathrm{D}-2)y$, 则原方程化为

$$(\mathrm{D}-1)u=x\mathrm{e}^{x+x^2},$$

这是一个关于 u 的一阶线性微分方程, 它可以用定理 8.1 求解, 它的一个特解为

$$u=\frac{1}{2}\,\mathrm{e}^{x+x^2}.$$

将此特解代入方程 $u=(\mathrm{D}-2)y$ 得

$$(\mathrm{D}-2)y=\frac{1}{2}\,\mathrm{e}^{x+x^2},$$

这是关于 y 的一阶线性微分方程. 由定理 8.1, 我们可得到此方程的一个 (满足 $y_1(0)=0$) 特解 $y_1(x)$:

$$y_1(x)=\frac{1}{2}\,\mathrm{e}^{2x}\int_0^x \mathrm{e}^{t^2-t}\,\mathrm{d}t.$$

尽管上述积分无法用初等函数表示, 我们仍然认为这个微分方程已经解出, 这是因为它的解已被表示为常用函数的积分. 方程 (8.25) 的通解为

$$y=c_1\mathrm{e}^{x}+c_2\mathrm{e}^{2x}+\frac{1}{2}\,\mathrm{e}^{2x}\int_0^x \mathrm{e}^{t^2-t}\,\mathrm{d}t.$$

□

8.14　求非齐次微分方程特解的零化子方法

本节介绍当微分方程 $L(y)=R$ 的系数为常数且函数 $R(x)$ 也是某个常系数齐次微分方程 $A(R)=0$ 的解时的一个求解方法. 这个方法的原理非常简单. 我们将算子 A 作用在微分方程 $L(y)=R$ 的两边, 由此得到一个新方程 $AL(y)=0$, 它必定被原来方程的所有解所满足. 由于 AL 是另一个常系数算子, 因此我们可以通过

求 AL 的相伴多项式的根来求出 AL 的零空间. 然后问题变成从这个零空间中选取满足 $L(y_1)=R$ 的一个特解 y_1. 我们用下面两个例子来说明解题过程.

例 1 求 4 阶微分方程

$$(\mathrm{D}^4-16)y=x^4+x+1$$

的一个特解.

解 因为方程右边是一个 4 次多项式, 所以 $\mathrm{D}^5(x^4+x+1)=0$. 因此, 原微分方程的任意一个解都是 9 阶方程

$$\mathrm{D}^5(\mathrm{D}^4-16)y=0 \tag{8.26}$$

的解. 由于相伴多项式的根为 $0,0,0,0,0,\pm2,\pm2\mathrm{i}$, 因此方程 (8.26) 的所有解都可从线性组合

$$y=c_1+c_2x+c_3x^2+c_4x^3+c_5x^4+c_6\mathrm{e}^{2x}+c_7\mathrm{e}^{-2x}+c_8\cos2x+c_9\sin2x$$

中找到. 我们要求出 c_i 使 $L(y)=x^4+x+1$, 此处 $L=\mathrm{D}^4-16$. 由于最后 4 项被 L 所零化, 因此可取 $c_6=c_7=c_8=c_9=0$, 然后求 $c_1,\cdots,c_5$ 使

$$L(c_1+c_2x+c_3x^2+c_4x^3+c_5x^4)=x^4+x+1.$$

换言之, 我们要找这样一个特解 y_1, 它是一个 4 次多项式, 并且满足条件 $L(y_1)=x^4+x+1$, 为避免引用下标以及使运算及表达式简洁, 我们令

$$16y_1=ax^4+bx^3+cx^2+dx+e.$$

将上式求导 4 次得 $16y_1^{(4)}=24a$, 因此 $y_1^{(4)}=3a/2$, 代入微分方程 $L(y_1)=x^4+x+1$, 我们得

$$\frac{3}{2}a-ax^4-bx^3-cx^2-dx-e=x^4+x+1.$$

比较上式两边 x 同次幂的系数, 我们得 $a=-1$, $b=c=0$, $d=-1$, $e=-\frac{5}{2}$, 因此特解 y_1 由下式给出:

$$y_1=-\frac{1}{16}x^4-\frac{1}{16}x-\frac{5}{32}.$$

直接代入微分方程可验证 y_1 确实是一个解. □

例 2 解微分方程 $y''-5y'+6=x\mathrm{e}^x$.

解 微分方程具有

$$L(y)=R \tag{8.27}$$

的形式, 其中 $R(x)=x\mathrm{e}^x$, $L=\mathrm{D}^2-5\mathrm{D}+6$. 相应的齐次方程为

$$(\mathrm{D}-2)(\mathrm{D}-3)y=0,$$

它有两个线性无关的特解 $u_1(x)=\mathrm{e}^{2x}$, $u_2(x)=\mathrm{e}^{3x}$. 下面求非齐次方程的一个特解 y_1. 我们把 $R(x)=x\mathrm{e}^x$ 看作齐次方程

$$(\mathrm{D}-1)^2y=0$$

的一个解. 于是将算子 $(\mathrm{D}-1)^2$ 作用于方程 (8.27) 的两边, 我们发现, 满足方程 (8.27) 的任意函数也满足方程

$$(D-1)^2(D-2)(D-3)y=0.$$

上述微分方程的相伴多项式的根为 1, 1, 2, 3, 因此, 这个微分方程的所有解都可从线性组合

$$y=ae^x+bxe^x+ce^{2x}+de^{3x}$$

中寻找, 其中 a,b,c,d 为常数. 我们希望选取的 a,b,c,d 使所得到的解 y_1 满足 $L(y_1)=xe^x$. 对任意 c 与 d, 都有 $L(ce^{2x}+de^{3x})=0$, 因此我们可令 $c=d=0$ 而选择 a 与 b 使 $L(ae^x+bxe^x)=xe^x$. 若令

$$y_1=ae^x+bxe^x,$$

则有

$$D(y_1)=(a+b)e^x+bxe^x,\quad D^2(y_1)=(a+2b)e^x+bxe^x,$$

因此方程 $(D^2-5D+6)y_1=xe^x$ 变成

$$(2a-3b)e^x+2bxe^x=xe^x.$$

从上式两边消去 e^x 再比较 x 的同次幂的系数, 我们得 $a=\frac{3}{4}$, $b=\frac{1}{2}$, 于是得

$$y_1=\frac{3}{4}e^x+\frac{1}{2}xe^x.$$

因此方程 $L(y)=R$ 的通解由下面的公式给出:

$$y=c_1e^{2x}+c_2e^{3x}+\frac{3}{4}e^x+\frac{1}{2}xe^x.$$

□

上述两个例子中所用的方法叫**零化子**(annihilator) 方法. 只要我们找到一个能将 R 零化的常系数算子 A, 就能使用这个方法. 不过就我们所掌握的关于常系数齐次线性微分方程的知识而言, 已经知道能被常系数算子零化的实值函数仅仅只有那些形如

$$x^{m-1}e^{\alpha x},\quad x^{m-1}e^{\alpha x}\cos\beta x,\quad x^{m-1}e^{\alpha x}\sin\beta x$$

的函数的线性组合的函数其中 m 是正整数, α 和 β 是实常数. 函数 $y=x^{m-1}e^{\alpha x}$ 是一个相伴多项式以 α 为 m 重根的微分方程的一个解, 因此这个函数被算子 $(D-\alpha)^m$ 零化. 函数 $y=x^{m-1}e^{\alpha x}\cos\beta x$ 与 $y=x^{m-1}e^{\alpha x}\sin\beta x$ 都是相伴多项式以复数 $\alpha\pm i\beta$ 为一对 m 重共轭复根的微分方程的解, 因此它们都被算子 $D^2-2\alpha D+(\alpha^2+\beta^2)$ 的 m 次幂所零化. 为便于读者参考, 我们将这些零化子连同它们的某些特殊情形在表 8.1 中列出.

表 8.1

函 数	零 化 子
$y=x^{m-1}$	D^m
$y=e^{\alpha x}$	$D-\alpha$
$y=x^{m-1}e^{\alpha x}$	$(D-\alpha)^m$
$y=\cos\beta x$ 或 $y=\sin\beta x$	$D^2+\beta^2$
$y=x^{m-1}\cos\beta x$ 或 $y=x^{m-1}\sin\beta x$	$(D^2+\beta^2)^m$
$y=e^{\alpha x}\cos\beta x$ 或 $y=e^{\alpha x}\sin\beta x$	$D^2-2\alpha D+(\alpha^2+\beta^2)$
$y=x^{m-1}e^{\alpha x}\cos\beta x$ 或 $y=x^{m-1}e^{\alpha x}\sin\beta x$	$[D^2-2\alpha D+(\alpha^2+\beta^2)]^m$

尽管零化子方法在能够应用的时候非常有效, 但它仅限于当方程右边的函数 R 有一个常系数零化子的情形才能使用. 如果 $R(x)$ 取 e^{x^2}, 或 $\ln x$, 或 $\tan x$ 等形式, 则此方法不再有用, 此时我们必须使用参数变易法或某些别的方法求方程的一个特解.

8.15 习　题

在习题 1～10 中, 求微分方程在区间 $(-\infty,+\infty)$ 上的通解.

1. $y''-y'=x^2$.
2. $y''-4y=e^{2x}$.
3. $y''+2y'=3xe^x$.
4. $y''+4y=\sin x$.
5. $y''-2y'+y=e^x+e^{2x}$.
6. $y'''-y'=e^x$.
7. $y'''-y'=e^x+e^{-x}$.
8. $y'''+3y''+3y'+y=xe^{-x}$.
9. $y''+y=xe^x\sin 2x$.
10. $y^{(4)}-y=x^2e^{-x}$.
11. 若 f 被一个常系数算子 A 零化, 又若常系数算子 B 零化 g, 证明: 乘积 AB 零化 $f+g$.
12. 设 A 为一个常系数算子, 其相伴多项式为 P_A.

    (a) 用零化子方法证明: 如果 α 不是多项式 P_A 的零点, 则微分方程 $A(y)=e^{\alpha x}$ 有一个形如
    $$y_1=\frac{e^{\alpha x}}{P_A(\alpha)}$$
    的特解.

    (b) 若 α 是 P_A 的单 (1 重) 零点, 证明: 方程 $A(y)=e^{\alpha x}$ 有特解
    $$y_1=\frac{xe^{\alpha x}}{P_A'(\alpha)}.$$

    (c) 将 (a) 与 (b) 的结果推广到 α 是 P_A 的 m 重零点的情形.
13. 给定两个常系数算子 A 与 B, 它们的相伴多项式没有公共零点, 令 $C=AB$.

    (a) 证明: 微分方程 $C(y)=0$ 的每一个解都可表成 $y=y_1+y_2$ 的形式, 其中 $A(y_1)=0$, $B(y_2)=0$.

    (b) 证明: (a) 中的函数 y_1 与 y_2 是唯一确定的. 即若给定 y 使 $C(y)=0$, 则存在唯一的一对函数 y_1 与 y_2 具有 (a) 中的性质.
14. 设 $L(y)=y''+ay'+by$, 其中 a 与 b 为常数. 再设 f 是方程 $L(y)=0$ 的满足条件 $f(0)=0$ 与 $f'(0)=1$ 的特解. 证明: 对任选的 c, 公式
    $$y_1(x)=\int_c^x f(x-t)R(t)\,dt$$
    都给出方程 $L(y)=R$ 的一个特解. 特别地, 当相伴多项式的两个根相等时, 设 $r_1=r_2=m$, 证明: 上述公式变化成
    $$y_1(x)=e^{mx}\int_c^x (x-t)e^{-mt}R(t)\,dt.$$
15. 令 Q 表示算子 "与 x 相乘", 即对 C^∞ 中每一个 y 及每个实数 x, 都有 $Q(y)(x)=xy(x)$. 再令 I 表示这样定义的恒等算子: 对 C^∞ 中每一个 y, 都有 $I(y)=y$.

    (a) 证明: $DQ-QD=I$.

    (b) 证明: D^2Q-QD^2 是一阶常系数算子, 给出将此算子表示为 D 的线性多项式的显式表示.

    (c) 证明: D^3Q-QD^3 是一个二阶常系数算子, 给出将此算子表示为 D 的二次多项式的显式表示.

    (d) 猜测关于 D^nQ-QD^n 的一个推广, 并用归纳法证明之.
16. 微分方程 $y''-3y'-4y=0$ 的一个解 $y=u(x)$ 的图像与微分方程 $y''+4y'-5y=0$ 的一个解

$y=v(x)$ 的图像交于原点. 如果这两条曲线在原点处有相同的斜率并且

$$\lim_{x\to\infty}\frac{[v(x)]^4}{u(x)}=\frac{5}{6},$$

试求出函数 u 与 v.

17. 微分方程 $y''-4y'+29y=0$ 的一个解 $y=u(x)$ 的图像与微分方程 $y''+4y'+13y=0$ 的一个解 $y=v(x)$ 的图像交于原点并且这两条曲线在原点处斜率相同. 若 $u'(\pi/2)=1$, 试确定 u 与 v.
18. 设微分方程 $y''+4xy'+Q(x)y=0$ 有两个如下形式的解: $y_1=u(x)$, $y_2=xu(x)$, 此处 $u(0)=1$. 试用 x 给出 $u(x)$ 与 $Q(x)$ 的显式表示.
19. 设 $L(y)=y''+P_1y'+P_2y$, 为用参数变易法解非齐次方程 $L(y)=R$, 我们需要知道对应齐次方程的两个线性无关的解. 本题指出, 如果 $L(y)=0$ 的一个解 u_1 已知, 并且 u_1 在一个区间 J 不为零, 则齐次方程的第二个解 u_2 由公式
$$u_2(x)=u_1(x)\int_c^x\frac{Q(t)}{[u_1(t)]^2}\,\mathrm{d}t$$
给出, 此处 $Q(x)=\mathrm{e}^{-A(x)}$, 其中 $A(x)=\int_a^x P_1(t)\,\mathrm{d}t$, a 与 c 为 J 中任意两点. 这两个解在 J 上线性无关.
    (a) 证明: 函数 u_2 确实满足 $L(y)=0$.
    (b) 对商 u_2/u_1 求导以证明 u_1 与 u_2 在 J 上线性无关.

    习题 19 可用来帮助解习题 20~23.
20. 若与非齐次微分方程 $xy''-2(x+1)y'+(x+2)y=x^3\mathrm{e}^{2x}$ 对应的齐次方程有一个形如 $y=\mathrm{e}^{mx}$ 的解, 试对 $x>0$ 求此非齐次方程的通解.
21. 如果已知当 $x>0$ 时微分方程 $4x^2y''+4xy'-y=0$ 有一个形如 $y=x^m$ 的特解, 求这个微分方程的通解.
22. 已知微分方程 $(2x-3x^2)y''+4y'+6xy=0$ 有一个解是 x 的多项式, 求这个方程的通解.
23. 若非齐次微分方程 $x^2(1-x)y''+2x(2-x)y'+2(1+x)y=x^2$ 对应的齐次方程有一个形如 $y=x^c$ 的解, 求这个非齐次微分方程的通解.
24. 设二阶微分方程 $y''+P(x)y'+Q(x)y=0$ 在正 x 轴上有解 $u(x)=x^3$, 两个线性无关解的 Wronski 行列式形如 $w(x)=cx^5$, 此处 c 为某个非零常数.
    (a) 给出 $P(x)$ 与 $Q(x)$ 的显式表示.
    (b) 求非齐次方程 $y''+P(x)y'+Q(x)y=x^2$ 在正 x 轴上的通解.

与线性微分方程相关的特征值问题

25. 设在 $(-\infty,+\infty)$ 上连续并且当 $t\longrightarrow-\infty$ 时 $f(t)\longrightarrow0$ 以及积分$\int_{-\infty}^x f(t)\,\mathrm{d}t$ 与 $\int_{-\infty}^x tf(t)\,\mathrm{d}t$ 对所有实数 x 都存在的所有函数 $f(x)$ 所组成的线性空间为 V, 对 $f\in V$, 由 $g(x)=x\int_{-\infty}^x f(t)\,\mathrm{d}t-\int_{-\infty}^x tf(t)\,\mathrm{d}t$ 定义 $g=T(f)$. 求出 T 的全体实特征值并且对每一个特征值 λ, 求相应的特征函数.
26. 令全体在区间 $[0,\pi]$ 上连续的实函数组成的线性空间为 V, 令在 $[0,\pi]$ 上有连续的二阶导数并且满足边界条件 $f(0)=f(\pi)=0$ 的全体函数 f 所组成的子空间为 S. 令 $T:S\longrightarrow V$ 为将各个 f 映为其二阶导数的线性变换, 即 $T(f)=f''$. 证明: T 的特征值是形如 $-n^2$ 的实数, 其中 $n=1,2,\cdots$, 它们对应的特征函数是 $f(t)=c_n\sin nt$, $c_n\neq0$.
27. 设所有在实轴上有任意阶导数的所有实函数 f 所组成的线性空间为 V, 再令所有满足边界条件 $f(0)=f(2\pi)$ 及 $f'(0)=f'(2\pi)=0$ 的全体 f 组成的子空间为 S. 令 $T:S\longrightarrow V$ 为二阶微分算子, 即 $T(f)=f''$. 求出 T 的全部 (无限多个) 特征值, 并且对每个特征值 λ 求出所有属于 λ 的特征函数.
28. 令全体处处都有任意阶导数并且满足条件 $f(0)=f''(0)$ 的实函数 f 组成的线性空间为 V. 令 $T:V\longrightarrow V$ 为二阶微分算子, $T(f)=f''$. 求 T 的全体特征值, 并且对每个特征值 λ, 求出所有属于 λ 的特征函数.

29. 令全体处处都有任意阶导数的实函数组成的线性空间为 V. 令 S 为由 V 中所有使得 $f(0) = 2f'(0)$ 的 f 组成的子空间. 再令 $T(f) = 4f''$. 求 T 的全体特征值, 并且对每个特征值 λ, 求 S 中所有属于 λ 的特征函数.
30. 令习题 29 中的线性空间为 V, 令 S 为由 V 中满足条件 $f(0) = f(1)$ 的函数 f 所组成的子空间. 再令 $T(f) = f''$. 求 T 的全体特征值, 并且对每一个特征值 λ, 求 S 中属于 λ 的全体特征函数.

第 9 章　在微分方程组理论中的应用

9.1　引　　言

尽管早在 17 世纪就开始了对微分方程的研究, 然而直到 19 世纪, 数学家们才意识到, 相对来说只有为数不多的微分方程才能用初等方法求解. Cauchy、Liouville 及其他一些数学家的工作, 表明了对某些特殊类型的微分方程建立保证其解存在的一般性定理的重要性. 第 8 章介绍了一个存在唯一性定理在线性微分方程研究中的应用. 本章将进一步发展指数矩阵等线性代数中的理论并将它们用于线性微分方程组的研究.

关于高阶微分方程的解的存在性理论可以化为一阶微分方程组来进行研究. 例如, 对二阶方程

$$y'' + 2ty' - y = \mathrm{e}^t, \tag{9.1}$$

可以通过引入两个未知函数 y_1 与 y_2 将它化成由两个一阶微分方程组成的方程组, 此处

$$y_1 = y, \quad y_2 = y_1'.$$

于是就有 $y_2' = y_1'' = y''$, 因此方程 (9.1) 可以写成如下的一阶微分方程组:

$$\begin{cases} y_1' = y_2, \\ y_2' = y_1 - 2ty_2 + \mathrm{e}^t. \end{cases} \tag{9.2}$$

由于这里的每一个方程都包含了两个未知函数, 因此无法用第 8 章中的方法来求解.

在本章中我们考虑由 n 个牵涉 n 个未知函数 $y_1, \cdots, y_n$ 的线性微分方程组成的方程组. 这些方程组有如下形式:

$$\begin{cases} y_1' = P_{11}(t)y_1 + P_{12}(t)y_2 + \cdots + P_{1n}(t)y_n + q_1(t), \\ \quad\vdots \\ y_n' = P_{n1}(t)y_1 + P_{n2}(t)y_2 + \cdots + P_{nn}(t)y_n + q_n(t). \end{cases} \tag{9.3}$$

在方程组 (9.3) 中出现的函数 P_{ik} 与 q_i 是定义在给定区间 J 上的已知函数, 函数 $y_1, \cdots, y_n$ 是有待确定的未知函数. 这类方程组叫做**一阶线性微分方程组**. 由于这个方程组中的各个方程一般来说都包含了多于 1 个的未知函数, 因此不可能将它们分开来求解.

所以要考虑微分方程组的理由之一是每个 n 阶线性微分方程都可以化为一阶线性微分方程组. 例如, 我们考察 n 阶线性微分方程

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = R(t), \tag{9.4}$$

其中系数 a_i 都是给定的函数. 为将此方程化为一阶线性微分方程组, 我们将 y 写成 y_1, 然后顺次对 y 的各阶导数引入一个新未知函数, 即我们令

$$y_1 = y, y_2 = y_1', y_3 = y_2', \cdots, y_n = y_{n-1}',$$

再将方程组 (9.4) 写成由 n 个方程组成的方程组:

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad \vdots \\ y_{n-1}' = y_n \\ y_n' = -a_n y_1 - a_{n-1} y_2 - \cdots - a_1 y_n + R(t). \end{cases} \tag{9.5}$$

采用向量与矩阵记号可以明显地简化关于线性微分方程组的讨论. 对一般的方程组 (9.3), 我们引入如下定义的向量值函数 $\boldsymbol{Y} = (y_1, \cdots, y_n)$, $\boldsymbol{Q} = (q_1, \cdots, q_n)$, $n \times n$ 矩阵值函数 $\boldsymbol{P} = [p_{ij}]$: 对任意 $t \in J$, 令

$$\boldsymbol{Y}(t) = (y_1(t), \cdots, y_n(t)), \boldsymbol{Q}(t) = (q_1(t), \cdots, q_n(t)), \boldsymbol{P}(t) = [p_{ij}(t)].$$

我们将向量 $\boldsymbol{Y}$ 与 $\boldsymbol{Q}$ 当作 $n \times 1$ 列矩阵并把方程组 (9.3) 表成如下简洁的形式:

$$\boldsymbol{Y}' = \boldsymbol{P}(t)\boldsymbol{Y} + \boldsymbol{Q}(t). \tag{9.6}$$

例如, 对方程组 (9.2), 我们有

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \ \boldsymbol{P}(t) = \begin{bmatrix} 0 & 1 \\ 1 & -2t \end{bmatrix}, \ \boldsymbol{Q}(t) = \begin{bmatrix} 0 \\ \mathrm{e}^t \end{bmatrix}.$$

对方程组 (9.5), 我们有

$$\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \ \boldsymbol{P}(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}, \ \boldsymbol{Q}(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ R(t) \end{bmatrix}.$$

对 $a \in J$ 和给定的一个 n 维常数向量 $\boldsymbol{B}$, 寻求满足方程组 (9.6) 与形如 $\boldsymbol{Y}(a) = \boldsymbol{B}$ 的初始条件的向量值函数 Y 的问题叫作方程组 (9.6) 的**初值问题**(initial-value problem). 无需多花任何力气, 就可将此问题推广到 $\boldsymbol{Q}(t), \boldsymbol{B}$ 和 $\boldsymbol{Y}(t)$ 都是 $n \times m$ 矩阵的情形.

在纯量 $(m = n = 1)$ 情形, 我们从定理 8.1 知道, 如果 P 与 Q 在 J 上连续, 则方程 (9.6) 的全部解都有下述显式表示

$$Y(x) = \mathrm{e}^{A(x)} Y(a) + \mathrm{e}^{A(x)} \int_a^x \mathrm{e}^{-A(t)} Q(t) \mathrm{d}t \tag{9.7}$$

给出, 其中 $A(x)=\int_a^x P(t)\mathrm{d}t$, a 为 J 中任意一点. 我们将证明, 这个公式能合适地推广到方程组的情形, 即当 $\boldsymbol{P}(t)$ 为 $n\times n$ 矩阵函数而 $\boldsymbol{Y}(t)$ 与 $\boldsymbol{Q}(t)$ 为 $n\times m$ 矩阵函数的情形. 不过在此之前, 必须定义矩阵的积分和指数矩阵, 因此我们简要地讨论矩阵函数的微积分.

9.2 矩阵函数的微积分

积分与微分的概念可以直接推广到矩阵函数的情形. 若 $\boldsymbol{P}(t)=[p_{ij}(t)]$ 为以函数为元素的任意 $n\times m$ 矩阵, 我们由

$$\int_a^b \boldsymbol{P}(t)\mathrm{d}t=\left[\int_a^b p_{ij}(t)\mathrm{d}t\right]$$

定义积分 $\int_a^b \boldsymbol{P}(t)\mathrm{d}t$, 即矩阵 $\boldsymbol{P}(t)$ 的积分等于以 $\boldsymbol{P}(t)$ 的各元素的积分为元素所组成的矩阵, 当然我们假定各元素在 $[a,b]$ 上都可积. 读者可以验证关于积分的线性性可以推广到矩阵函数的情形.

矩阵函数的连续性与可微性等概念也可通过元素来定义. 如果每个元素 p_{ij} 都在 t 点连续, 则称矩阵函数 $\boldsymbol{P}=[p_{ij}]$ 在 t 点连续. 当所有导数 $p'_{ij}(t)$ 都存在时, 可以通过各元素的导数来定义矩阵函数的导数 $\boldsymbol{P}'$ 如下:

$$\boldsymbol{P}'(t)=[p'_{ij}(t)].$$

容易验证关于函数和与积的基本微分法则对矩阵函数也成立. 例如, 若 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 都是可微矩阵函数, 则当 $\boldsymbol{P}$ 与 $\boldsymbol{Q}$ 为行数与列数相同的矩阵时, 我们有

$$(\boldsymbol{P}+\boldsymbol{Q})'=\boldsymbol{P}'+\boldsymbol{Q}';$$

而且当乘积 $\boldsymbol{PQ}$ 有意义时, 我们也有

$$(\boldsymbol{PQ})'=\boldsymbol{PQ}'+\boldsymbol{P}'\boldsymbol{Q}.$$

关于复合函数的求导法则也成立. 即若 $\boldsymbol{F}(t)=\boldsymbol{P}[g(t)]$, 则当 $\boldsymbol{P}$ 是可微矩阵函数且 g 为可微纯量函数时, 我们有 $\boldsymbol{F}'(t)=g'(t)\boldsymbol{P}'[g(t)]$. 微积分第一基本定理与第二基本定理以及零导数定理对矩阵函数同样成立. 这些性质的证明作为 9.4 节中的一组习题留给读者.

然而指数矩阵的定义并不如此简单, 需要作进一步的准备, 这将在下面几节中讨论.

9.3 矩阵幂级数 · 矩阵的范数

设 $\boldsymbol{A}=[a_{ij}]$ 为 $n\times n$ 实矩阵或复矩阵, 我们希望以这样一种方式来定义指数 $\mathrm{e}^{\boldsymbol{A}}$, 使得它具有关于通常实指数或复值指数的某些基本性质. 特别地, 我们希望这样的定义保持如下形式的**指数律**:

$$e^{tA}e^{sA} = e^{(s+t)A}, \quad 对所有实数 s 与 t 都成立, \tag{9.8}$$

以及关系式

$$e^{O} = I, \tag{9.9}$$

其中 O 与 I 分别为 $n \times n$ 零矩阵与单位矩阵. 似乎我们可以很自然地将 e^{A} 定义为以各个元素 a_{ij} 作指数所得到的矩阵 $[e^{a_{ij}}]$, 但这在实际上是不能接受的, 因为这样的定义既不满足性质 (9.8) 也不满足性质 (9.9). 因而我们不采用这个定义, 而是用幂级数展开式

$$e^{A} = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

来定义 e^{A}. 我们知道, 上述公式当 A 为实数时成立 (当 A 为复数时也成立), 我们将证明当 A 为矩阵时, 由上式可推出性质 (9.8) 与 (9.9). 在此之前, 我们还需说明收敛矩阵级数是什么意思.

定义 9.1 (收敛矩阵级数) 给定实或复 $m \times n$ 矩阵的无穷序列 $\{C_k\}$, 令 $c_{ij}^{(k)}$ 表示 C_k 的 (i, j)- 元. 若全部 mn 个级数

$$\sum_{k=1}^{\infty} c_{ij}^{(k)}, \quad (i = 1, \cdots, m; j = 1, \cdots, n) \tag{9.10}$$

都收敛, 则称矩阵级数 $\sum_{k=1}^{\infty} C_k$ 收敛, 并且定义它的和是以级数 (9.10) 作为 (i, j)- 元的 $m \times n$ 矩阵.

关于矩阵级数收敛性的一个简单而且很有用的判别法可以用矩阵的范数来给出. 矩阵的范数可以看作是数的绝对值的一个推广.

定义 9.2 (矩阵的范数) 设 $A = [a_{ij}]$ 为一个 $m \times n$ 实矩阵或复矩阵, 令 $\|A\|$ 表示 A 的所有元素的绝对值之和, 即

$$\|A\| = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|, \tag{9.11}$$

则称 $\|A\|$ 为矩阵 A 的**范数**(norm).

有时也用别的方式定义范数, 我们采用上述定义是因为由此可以很容易导出下列性质.

定理 9.1 (范数的基本性质) 设 A 与 B 为矩阵, c 为任意实数或复数, 则我们有

$$\|A + B\| \leqslant \|A\| + \|B\|, \quad \|AB\| \leqslant \|A\|\|B\|, \quad \|cA\| = |c|\|A\|.$$

证明 我们只给出 $\|AB\| \leqslant \|A\|\|B\|$ 的证明, 另外两个性质的证明更为简单, 我们将它们作为习题留给读者.

设 A 为 $m \times n$ 矩阵, B 为 $n \times p$ 矩阵. 若 $A = [a_{ik}], B = [b_{kj}]$, 则 $AB = \left[\sum_{k=1}^{n} a_{ik}b_{kj}\right]$, 因此由等式 (9.11) 得

$$\|AB\| = \sum_{i=1}^{m}\sum_{j=1}^{p}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \leqslant \sum_{i=1}^{m}\sum_{k=1}^{n}|a_{ik}|\sum_{j=1}^{p}|b_{kj}|$$

$$\leqslant \sum_{i=1}^{m}\sum_{k=1}^{n}|a_{ik}|\|B\| = \|A\|\|B\|. \qquad \square$$

注意到在 $B = A$ 的特殊情形 (此时两者都是 $n\times n$ 矩阵), 关于 $\|AB\|$ 的不等式变成 $\|A^2\| \leqslant \|A\|^2$. 由对 k 的归纳法又得, 对任意 $n\times n$ 矩阵 A 都有

$$\|A^k\| \leqslant \|A\|^k, k = 2, 3, \cdots .$$

这些不等式将用于关于指数矩阵的讨论.

下述定理给出关于矩阵级数收敛性的一个充分条件.

定理 9.2 (矩阵级数收敛性的判别法) 设 $\{C_k\}$是一个 $n\times n$ 矩阵的序列使得范数级数 $\sum\limits_{k=1}^{\infty}\|C_k\|$ 收敛, 则矩阵级数 $\sum\limits_{k=1}^{\infty} C_k$ 也收敛.

证明 令 $c_{ij}^{(k)}$ 表示 C_k 的 (i,j)- 元. 由于 $|c_{ij}^{(k)}| \leqslant \|C_k\|$, 因此由 $\sum\limits_{k=1}^{\infty}\|C_k\|$ 的收敛性即得各级数 $\sum\limits_{k=1}^{\infty} c_{ij}^{(k)}$ 的绝对收敛性, 从而各个级数 $\sum\limits_{k=1}^{\infty} c_{ij}^{(k)}$ 都收敛, 即矩阵级数 $\sum\limits_{k=1}^{\infty} C_k$ 收敛. $\square$

9.4 习 题

1. 证明: 积分的线性性对矩阵函数的积分也成立.
2. 设 P 与 Q 都可微, 验证关于矩阵函数的下列微分法则. 在 (a) 中, P 与 Q 必须行数相同且列数也相同才能使 $P+Q$ 有意义. 在 (b) 与 (d) 中, P 的列数必须和 Q 的行数相同才能使它们的乘积有意义. 在 (c) 与 (d) 中假定 Q 是可逆矩阵.

   (a) $(P+Q)' = P' + Q'$. (b) $(PQ)' = PQ' + P'Q$.

   (c) $(Q^{-1})' = -Q^{-1}Q'Q^{-1}$. (d) $(PQ^{-1})' = -PQ^{-1}Q'Q^{-1} + P'Q^{-1}$.
3. (a) 设 P 为可微矩阵函数, 证明: P^2 与 P^3 的导数由下列公式给出

   $$(P^2)' = PP' + P'P, \quad (P^3)' = P^2P' + PP'P + P'P^2.$$

   (b) 猜测出关于 P^k 的导数的一般公式, 然后对 k 作归纳法来证明这个公式.
4. 设 P 为可微矩阵函数, g 为可微纯量函数, g 的值域是 P 的定义域的一个子集. 令 F 为由 $F(t) = P[g(t)]$ 所定义的复合函数, 证明复合函数求导的**链式法则**: $F'(t) = g'(t)P'[g(t)]$.
5. 证明关于矩阵函数的**零导数定理**: 若对某个开区间 (a,b) 内的所有 t 都有 $P'(t) = O$, 则矩阵函数 $P(t)$ 在 (a,b) 上是一个常数矩阵.
6. 将微积分第一基本定理和第二基本定理推广到矩阵函数并给出证明.
7. 叙述并证明以矩阵函数作为被积函数的一个分部积分公式.
8. 证明矩阵范数的下列性质:

   $$\|A+B\| \leqslant \|A\| + \|B\|, \quad \|cA\| = |c|\|A\|.$$
9. 设矩阵函数 P 在区间 $[a,b]$ 上可积, 证明:

$$\left\|\int_a^b \boldsymbol{P}(t)\mathrm{d}t\right\| \leqslant \int_a^b \|\boldsymbol{P}(t)\|\mathrm{d}t.$$

10. 设 $\boldsymbol{D}$ 为一个 $n\times n$ 对角矩阵, $\boldsymbol{D}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$, 证明: 矩阵级数 $\sum\limits_{k=0}^{\infty}\frac{\boldsymbol{D}^k}{k!}$ 收敛并且也是一个对角矩阵:

$$\sum_{k=0}^{\infty}\frac{\boldsymbol{D}^k}{k!}=\operatorname{diag}\left(\mathrm{e}^{\lambda_1},\cdots,\mathrm{e}^{\lambda_n}\right).$$

(对应于 $k=0$ 的项约定为单位矩阵 $\boldsymbol{I}$.)

11. 设 $\boldsymbol{D}$ 为一个 $n\times n$ 单位矩阵, $\boldsymbol{D}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. 若矩阵级数 $\sum\limits_{k=0}^{\infty}c_k\boldsymbol{D}^k$ 收敛, 证明:

$$\sum_{k=0}^{\infty}c_k\boldsymbol{D}^k=\operatorname{diag}\left(\sum_{k=0}^{\infty}c_k\lambda_1^k,\cdots,\sum_{k=0}^{\infty}c_k\lambda_n^k\right).$$

12. 设矩阵级数 $\sum\limits_{k=1}^{\infty}\boldsymbol{C}_k$ 收敛, 其中各 $\boldsymbol{C}_k$ 为 $n\times n$ 矩阵. 证明: 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为使乘积 $\boldsymbol{AC}_k\boldsymbol{B}$ 有意义的矩阵, 则矩阵级数 $\sum\limits_{k=1}^{\infty}(\boldsymbol{AC}_k\boldsymbol{B})$ 也收敛并且其和为矩阵 $\boldsymbol{A}\left(\sum\limits_{k=1}^{\infty}\boldsymbol{C}_k\right)\boldsymbol{B}$.

9.5 指数矩阵

利用定理 9.2 的判别法容易证明矩阵级数

$$\sum_{k=0}^{\infty}\frac{\boldsymbol{A}^k}{k!} \tag{9.12}$$

对任意实矩阵或复矩阵都收敛 (对应于 $k=0$ 的项为单位矩阵 $\boldsymbol{I}$). 各项的范数满足不等式

$$\left\|\frac{\boldsymbol{A}^k}{k!}\right\| \leqslant \frac{\|\boldsymbol{A}\|^k}{k!}=\frac{a^k}{k!},$$

此处 $a=\|\boldsymbol{A}\|$. 由于级数 $\sum\limits_{k=0}^{\infty}\frac{a^k}{k!}$ 对任意实数 a 都收敛, 因此由定理 9.2 即得矩阵级数 (9.12) 对任意方阵 $\boldsymbol{A}$ 都收敛. □

定义 9.3 (指数矩阵) 对任意 $n\times n$ 实矩阵或复矩阵 $\boldsymbol{A}$, 由收敛级数 (9.12) 所给出的 $n\times n$ 矩阵

$$\sum_{k=0}^{\infty}\frac{\boldsymbol{A}^k}{k!}$$

叫作**指数矩阵**(exponential matrix) 并记作 $\mathrm{e}^{\boldsymbol{A}}$. □

由上述定义有 $\mathrm{e}^{\boldsymbol{O}}=\boldsymbol{I}$, 此处 $\boldsymbol{O}$ 为零矩阵. 关于指数矩阵的进一步性质将借助于微分方程展开讨论.

9.6 e^{tA} 所满足的微分方程

设 t 为实数, $\boldsymbol{A}$ 为 $n\times n$ 矩阵, 令 $\boldsymbol{E}(t)$ 为由

$$\boldsymbol{E}(t) = \mathrm{e}^{t\boldsymbol{A}}$$

所给出的 $n \times n$ 矩阵. 我们将 $\boldsymbol{A}$ 固定而把 $\boldsymbol{E}(t)$ 看作 t 的函数. 我们先给出 E 所满足的一个微分方程.

定理 9.3　对任意实数 t, 由 $\boldsymbol{E}(t) = \mathrm{e}^{t\boldsymbol{A}}$ 定义的矩阵函数 $\boldsymbol{E}$ 都满足矩阵微分方程

$$\boldsymbol{E}'(t) = \boldsymbol{E}(t)\boldsymbol{A} = \boldsymbol{A}\boldsymbol{E}(t).$$

证明　由指数矩阵的定义有

$$\boldsymbol{E}(t) = \sum_{k=0}^{\infty} \frac{(t\boldsymbol{A})^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k \boldsymbol{A}^k}{k!}.$$

令 $c_{ij}^{(k)}$ 表示 $\boldsymbol{A}^k$ 的 (i,j)- 元, 则 $\frac{t^k\boldsymbol{A}^k}{k!}$ 的 (i,j)- 元是 $\frac{t^k c_{ij}^{(k)}}{k!}$. 因此由矩阵级数的定义得

$$\sum_{k=0}^{\infty} \frac{t^k \boldsymbol{A}^k}{k!} = \left[\sum_{k=0}^{\infty} \frac{t^k}{k!} c_{ij}^{(k)}\right]. \tag{9.13}$$

等式 (9.13) 右边的各元素都是 t 的幂级数, 它们对所有 t 都收敛, 因此对所有 t, 其导数都存在并且由求导后的级数

$$\sum_{k=1}^{\infty} \frac{k t^{k-1}}{k!} c_{ij}^{(k)} = \sum_{k=0}^{\infty} \frac{t^k}{k!} c_{ij}^{(k+1)}$$

给出. 由此证明了导数 $\boldsymbol{E}'(t)$ 的存在性并且由矩阵级数

$$\boldsymbol{E}'(t) = \sum_{k=0}^{\infty} \frac{t^k \boldsymbol{A}^{k+1}}{k!} = \left(\sum_{k=0}^{\infty} \frac{t^k \boldsymbol{A}^k}{k!}\right) \boldsymbol{A} = \boldsymbol{E}(t)\boldsymbol{A}$$

给出. 在上面最后一式中我们用到了性质 $\boldsymbol{A}^{k+1} = \boldsymbol{A}^k\boldsymbol{A}$. 由于 $\boldsymbol{A}$ 与 $\boldsymbol{A}^k$ 交换, 因此也有 $\boldsymbol{A}^{k+1} = \boldsymbol{A}\boldsymbol{A}^k$, 由此得 $\boldsymbol{E}'(t) = \boldsymbol{A}\boldsymbol{E}(t)$. 即得定理. □

注　上述证明也指出了 $\mathrm{e}^{t\boldsymbol{A}}$ 与 $\boldsymbol{A}$ 可交换.

9.7　矩阵微分方程 $F'(t) = AF(t)$ 的解的唯一性定理

在本节中, 我们要证明一个刻划矩阵微分方程 $\boldsymbol{F}'(t) = \boldsymbol{A}\boldsymbol{F}(t)$ 的所有解的特征的唯一性定理. 这个定理的证明用到下述定理.

定理 9.4 (e^{tA} 的非奇异性)　对任意 $n \times n$ 矩阵 $\boldsymbol{A}$ 和任意纯量 t, 我们都有

$$\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{-t\boldsymbol{A}} = \boldsymbol{I}, \tag{9.14}$$

因此 $\mathrm{e}^{t\boldsymbol{A}}$ 是非奇异矩阵而且它的逆矩阵是 $\mathrm{e}^{-t\boldsymbol{A}}$.

证明　令 F 为由

$$\boldsymbol{F}(t) = \mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{-t\boldsymbol{A}}$$

所定义的矩阵函数, 其中 t 为任意实数. 我们通过证明导数 $\boldsymbol{F}'(t)$ 为零矩阵来证明 $\boldsymbol{F}(t)$ 是单位矩阵. 对 $\boldsymbol{F}$ 求导, 由于 $\boldsymbol{A}$ 与 $\mathrm{e}^{t\boldsymbol{A}}$ 可交换, 因此由乘积的求导法则及定理 9.3 我们得

$$\begin{aligned}\boldsymbol{F}'(t) &= \mathrm{e}^{t\boldsymbol{A}}(\mathrm{e}^{-t\boldsymbol{A}})' + (\mathrm{e}^{t\boldsymbol{A}})'\mathrm{e}^{-t\boldsymbol{A}} \\ &= \mathrm{e}^{t\boldsymbol{A}}(-\boldsymbol{A}\mathrm{e}^{-t\boldsymbol{A}}) + \boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{-t\boldsymbol{A}} \\ &= -\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{-t\boldsymbol{A}} + \boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{-t\boldsymbol{A}} \\ &= \boldsymbol{O}.\end{aligned}$$

于是由零导数定理得 $\boldsymbol{F}$ 是一个常数矩阵. 又由于 $\boldsymbol{F}(0) = \mathrm{e}^{\boldsymbol{O}}\mathrm{e}^{-\boldsymbol{O}} = \boldsymbol{I}$, 因此对所有 t 都有 $\boldsymbol{F}(t) = \boldsymbol{I}$, 即得等式 (9.14). □

定理 9.5 (唯一性定理) 设 $\boldsymbol{A}$ 为给定的 $n \times n$ 常数矩阵, $\boldsymbol{B}$ 为给定的 $n \times m$ 常数矩阵, 则微分方程

$$\boldsymbol{F}'(t) = \boldsymbol{A}\boldsymbol{F}(t)$$

的对所有实数 t 都满足初始条件

$$\boldsymbol{F}(0) = \boldsymbol{B}$$

的解为 $n \times m$ 矩阵函数 $\boldsymbol{F}$:

$$\boldsymbol{F}(t) = \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}. \tag{9.15}$$

证明 首先我们易知 $\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}$ 是一个解. 现在设 $\boldsymbol{F}$ 为任意一个解, 考虑矩阵函数

$$\boldsymbol{G}(t) = \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{F}(t).$$

两边求导数得

$$\boldsymbol{G}'(t) = \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{F}'(t) - \boldsymbol{A}\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{F}(t) = \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{A}\boldsymbol{F}(t) - \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{A}\boldsymbol{F}(t) = \boldsymbol{O},$$

从而 $\boldsymbol{G}(t)$ 是一个常数矩阵, 因此

$$\boldsymbol{G}(t) = \boldsymbol{G}(0) = \boldsymbol{F}(0) = \boldsymbol{B},$$

即 $\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{F}(t) = \boldsymbol{B}$, 两边同乘以 $\mathrm{e}^{t\boldsymbol{A}}$, 由恒等式 (9.14) 即得

$$\boldsymbol{F}(t) = \mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}.$$

因此解唯一, 即得定理. □

注 同理可证 $\boldsymbol{F}(t) = \boldsymbol{B}\mathrm{e}^{t\boldsymbol{A}}$ 是初值问题

$$\boldsymbol{F}'(t) = \boldsymbol{F}(t)\boldsymbol{A}, \ \boldsymbol{F}(0) = \boldsymbol{B}$$

的唯一解, 此处 $\boldsymbol{F}(t)$ 与 $\boldsymbol{B}$ 为 $m \times n$ 矩阵而 $\boldsymbol{A}$ 为 $n \times n$ 矩阵.

9.8 关于指数矩阵的指数律

指数律 $\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}$ 并不是对所有矩阵指数都成立的, 9.11 节习题 13 给出了这样的一个反例. 然而不难证明, 当矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可交换时指数律一定成立.

定理 9.6 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个可交换的 $n \times n$ 矩阵: $\boldsymbol{AB} = \boldsymbol{BA}$, 则

$$\mathrm{e}^{\boldsymbol{A}}\mathrm{e}^{\boldsymbol{B}} = \mathrm{e}^{\boldsymbol{A}+\boldsymbol{B}}. \tag{9.16}$$

证明 由 $\boldsymbol{AB} = \boldsymbol{BA}$ 得

$$\boldsymbol{A}^2\boldsymbol{B} = \boldsymbol{A}(\boldsymbol{BA}) = (\boldsymbol{AB})\boldsymbol{A} = (\boldsymbol{BA})\boldsymbol{A} = \boldsymbol{BA}^2,$$

因此 $\boldsymbol{B}$ 与 $\boldsymbol{A}^2$ 可交换. 由归纳法可得 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 的任意正整数幂都可交换. 通过将 $\mathrm{e}^{t\boldsymbol{A}}$ 表示成幂级数形式可以看出, 对任意实数 t, $\boldsymbol{B}$ 与 $\mathrm{e}^{t\boldsymbol{A}}$ 也可交换.

现在设 F 为由

$$\boldsymbol{F}(t)=\mathrm{e}^{t(\boldsymbol{A}+\boldsymbol{B})}-\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{t\boldsymbol{B}}$$

所定义的矩阵函数, 我们要证明, 对所有 t, $\boldsymbol{F}(t)$ 都等于 $\boldsymbol{O}$. 对 $\boldsymbol{F}(t)$ 求导并利用 $\boldsymbol{B}$ 与 $\mathrm{e}^{t\boldsymbol{A}}$ 可交换这一性质, 我们有

$$\begin{aligned}\boldsymbol{F}'(t)&=(\boldsymbol{A}+\boldsymbol{B})\mathrm{e}^{t(\boldsymbol{A}+\boldsymbol{B})}-\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{t\boldsymbol{B}}-\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}\mathrm{e}^{t\boldsymbol{B}}\\&=(\boldsymbol{A}+\boldsymbol{B})\mathrm{e}^{t(\boldsymbol{A}+\boldsymbol{B})}-(\boldsymbol{A}+\boldsymbol{B})\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{t\boldsymbol{B}}\\&=(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{F}(t).\end{aligned}$$

从而由唯一性定理得

$$\boldsymbol{F}(t)=\mathrm{e}^{t(\boldsymbol{A}+\boldsymbol{B})}\boldsymbol{F}(0).$$

而由于 $\boldsymbol{F}(0)=\boldsymbol{O}$, 因此对所有 t 都有 $\boldsymbol{F}(t)=\boldsymbol{O}$, 于是得

$$\mathrm{e}^{t(\boldsymbol{A}+\boldsymbol{B})}=\mathrm{e}^{t\boldsymbol{A}}\mathrm{e}^{t\boldsymbol{B}}.$$

当 $t=1$ 时即得等式 (9.16). □

例　矩阵 $s\boldsymbol{A}$ 与 $t\boldsymbol{A}$ 对所有纯量 s 与 t 都可交换, 因此我们有

$$\mathrm{e}^{s\boldsymbol{A}}\mathrm{e}^{t\boldsymbol{A}}=\mathrm{e}^{(s+t)\boldsymbol{A}}.$$

9.9　常系数齐次线性微分方程组的存在唯一性定理

设 $\boldsymbol{A}$ 为 $n\times n$ 常数矩阵, $\boldsymbol{Y}$ 为 $n\times m$ 矩阵函数, 则矩阵微分方程 $\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t)$ 叫做**常系数线性 (微分) 方程组**. 我们将利用指数矩阵给出这类微分方程组的解的显式表示.

定理 9.7　*设 $\boldsymbol{A}$ 为给定的 $n\times n$ 常数矩阵, $\boldsymbol{B}$ 为给定的 $n\times m$ 常数矩阵, 则初值问题*

$$\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t),\ \boldsymbol{Y}(0)=\boldsymbol{B}\tag{9.17}$$

在区间 $-\infty<t<+\infty$ 内有唯一的 $n\times m$ 矩阵解且其解由公式

$$\boldsymbol{Y}(t)=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}\tag{9.18}$$

给出. 更一般地, 初值问题

$$\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t),\ \boldsymbol{Y}(a)=\boldsymbol{B}$$

的解为

$$\boldsymbol{Y}(t)=\mathrm{e}^{(t-a)\boldsymbol{A}}\boldsymbol{B}.$$

证明　对式 (9.18) 两边求导得 $\boldsymbol{Y}'(t)=\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}=\boldsymbol{A}\boldsymbol{Y}(t)$. 由于 $\boldsymbol{Y}(0)=\boldsymbol{B}$, 因此这是初值问题 (9.17) 的一个解. 为证明这是唯一解, 我们采用定理 9.5 的证法. 设 $\boldsymbol{Z}(t)$ 为另一满足 $\boldsymbol{Z}'(t)=\boldsymbol{A}\boldsymbol{Z}(t)$ 且 $\boldsymbol{Z}(0)=\boldsymbol{B}$ 的矩阵函数, 令 $\boldsymbol{G}(t)=\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Z}(t)$, 则容易验证 $\boldsymbol{G}'(t)=\boldsymbol{O}$, 因此 $\boldsymbol{G}(t)=\boldsymbol{G}(0)=\boldsymbol{Z}(0)=\boldsymbol{B}$, 即 $\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Z}(t)=\boldsymbol{B}$, 于是 $\boldsymbol{Z}(t)=\mathrm{e}^{t\boldsymbol{A}}\boldsymbol{B}=\boldsymbol{Y}(t)$. 对 $\boldsymbol{Y}(a)=\boldsymbol{B}$ 的较一般情形可以用完全相同的方法处理. □

9.10 在特殊情形下 e^{tA} 的计算

虽然定理 9.7 给出了常系数齐次线性微分方程组解的显式表示, 但是还存在实际计算指数矩阵 $\mathrm{e}^{t\boldsymbol{A}}$ 的问题. 如果想从级数定义出发计算 $\mathrm{e}^{t\boldsymbol{A}}$, 则我们就得计算所有的幂 $\boldsymbol{A}^k$, $k=0,1,2,\cdots$. 然后还要计算各个级数 $\sum\limits_{k=0}^{\infty}\frac{t^k c_{ij}^{(k)}}{k!}$ 的和, 此处 $c_{ij}^{(k)}$ 是 $\boldsymbol{A}^k$ 的 (i,j)- 元. 一般来说, 这是一个没有希望完成的任务, 除非矩阵 $\boldsymbol{A}$ 的各次幂能很容易地求出. 例如, 若 $\boldsymbol{A}$ 是对角矩阵, 不妨设

$$\boldsymbol{A}=\mathrm{diag}\,(\lambda_1,\cdots,\lambda_n),$$

则 $\boldsymbol{A}$ 的任意次幂也是对角矩阵, 事实上, $\boldsymbol{A}^k=\mathrm{diag}\,(\lambda_1^k,\cdots,\lambda_n^k)$. 因而在这一情形 $\mathrm{e}^{t\boldsymbol{A}}$ 是由

$$\mathrm{e}^{tA}=\mathrm{diag}\,\left(\sum_{k=0}^{\infty}\frac{t^k}{k!}\lambda_1^k,\cdots,\sum_{k=0}^{\infty}\frac{t^k}{k!}\lambda_n^k\right)=\mathrm{diag}\,(\mathrm{e}^{t\lambda_1},\cdots,\mathrm{e}^{t\lambda_n})$$

确定的对角矩阵. 另一种简单情形是 $\boldsymbol{A}$ 为一个可以对角化的矩阵. 例如, 若存在非奇异矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ 是对角矩阵, 设 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{D}$, 则 $\boldsymbol{A}=\boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^{-1}$, 由此可得

$$\boldsymbol{A}^2=(\boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^{-1})(\boldsymbol{C}\boldsymbol{D}\boldsymbol{C}^{-1})=\boldsymbol{C}\boldsymbol{D}^2\boldsymbol{C}^{-1},$$

以及对任意正整数 k, 都有

$$\boldsymbol{A}^k=\boldsymbol{C}\boldsymbol{D}^k\boldsymbol{C}^{-1}.$$

从而在这种情形下有

$$\mathrm{e}^{t\boldsymbol{A}}=\sum_{k=0}^{\infty}\frac{t^k\boldsymbol{A}^k}{k!}=\sum_{k=0}^{\infty}\frac{t^k}{k!}\boldsymbol{C}\boldsymbol{D}^k\boldsymbol{C}^{-1}=\boldsymbol{C}\left(\sum_{k=0}^{\infty}\frac{t^k\boldsymbol{D}^k}{k!}\right)\boldsymbol{C}^{-1}=\boldsymbol{C}\mathrm{e}^{t\boldsymbol{D}}\boldsymbol{C}^{-1}.$$

对角矩阵 $\boldsymbol{D}$ 与 $\boldsymbol{A}$ 有相同的特征值, 设为 $\lambda_1,\cdots,\lambda_n$, 则

$$\mathrm{e}^{t\boldsymbol{D}}=\mathrm{diag}\,\left(\mathrm{e}^{t\lambda_1},\cdots,\mathrm{e}^{t\lambda_n}\right),$$

因此 $\boldsymbol{C}\mathrm{e}^{t\boldsymbol{D}}\boldsymbol{C}^{-1}$ 是形如

$$\mathrm{e}^{t\lambda_1}L_1(\boldsymbol{A})+\mathrm{e}^{t\lambda_2}L_2(\boldsymbol{A})+\cdots+\mathrm{e}^{t\lambda_n}L_n(\boldsymbol{A})$$

的一个线性组合, 其中各**乘子**(multiplier)$L_1(\boldsymbol{A}),L_2(\boldsymbol{A}),\cdots,L_n(\boldsymbol{A})$ 都是依赖于 $\boldsymbol{C}$ 从而也依赖于 $\boldsymbol{A}$ 的 $n\times n$ 矩阵. 换言之, 我们已经证明了下述定理中的展开式 (9.19).

定理 9.8 若 $n\times n$ 矩阵可对角化, 并且 $\boldsymbol{A}$ 的特征值为 $\lambda_1,\cdots,\lambda_n$, 则

$$\mathrm{e}^{t\boldsymbol{A}}=\sum_{k=1}^{n}\mathrm{e}^{t\lambda_k}L_k(\boldsymbol{A}),\tag{9.19}$$

此处各乘子 $L_k(\boldsymbol{A})$ 都是仅依赖于 $\boldsymbol{A}$ 的 $n\times n$ 矩阵. 由此又有

$$\boldsymbol{A}^r = \sum_{k=1}^{n} \lambda_k^r L_k(\boldsymbol{A}), \qquad \text{对每一个整数 } r \geqslant 0, \tag{9.20}$$

$$p(\boldsymbol{A}) = \sum_{k=1}^{n} p(\lambda_k) L_k(\boldsymbol{A}), \qquad \text{对任意多项式 } p(\lambda). \tag{9.21}$$

证明 在展开式 (9.19) 中取 $t=0$ 得

$$\boldsymbol{I} = \sum_{k=1}^{n} L_k(\boldsymbol{A}).$$

对展开式 (9.19) 两边求导得

$$\boldsymbol{A}\mathrm{e}^{t\boldsymbol{A}} = \sum_{k=1}^{n} \lambda_k \mathrm{e}^{t\lambda_k} L_k(\boldsymbol{A}),$$

再令 $t=0$, 则得

$$\boldsymbol{A} = \sum_{k=1}^{n} \lambda_k L_k(\boldsymbol{A}).$$

继续这一程序, 对展开式 (9.19) 左边每求导一次都将使左边新增加一个因子 $\boldsymbol{A}$, 同时使右边和式中的第 k 项增加一个新因子 λ_k. 经过 r 次求导然后取 $t=0$, 就得到展开式 (9.20).

现在令$p(\lambda) = \sum\limits_{r=0}^{m} c_r \lambda^r$ 为 λ 的任意多项式, 对 $r=0,1,\cdots,m$, 将展开式 (9.20) 两边各项乘以 c_r, 然后相加即得展开式 (9.21). □

如下述定理所示, 适当选择多项式 $p(\lambda)$, 我们可得到展开式 (9.19) 中矩阵乘子 $L_k(A)$ 的显式表示.

定理 9.9 若 $n\times n$ 矩阵 $\boldsymbol{A}$ 有 n 个不同的特征值 $\lambda_1,\cdots,\lambda_n$, 则

$$\mathrm{e}^{t\boldsymbol{A}} = \sum_{k=1}^{n} \mathrm{e}^{t\lambda_k} L_k(\boldsymbol{A}), \tag{9.22}$$

此处 $L_k(A)$ 是由公式

$$L_k(\boldsymbol{A}) = \prod_{\substack{j=1\\ j\neq k}}^{n} \frac{\boldsymbol{A} - \lambda_j \boldsymbol{I}}{\lambda_k - \lambda_j}, \quad k = 1,2,\cdots,n \tag{9.23}$$

给出的 $\boldsymbol{A}$ 的 $n-1$ 次多项式.

证明 由于 $\boldsymbol{A}$ 有 n 个不同的特征值, 因此 $\boldsymbol{A}$ 可以对角化. 于是由定理 9.8, 对任意多项式 $p(\lambda)$, 等式 (9.19) 与等式 (9.21) 都成立. 如果取 $p(\lambda)$ 为 $\boldsymbol{A}$ 的特征多项式, 设

$$p(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n), \tag{9.24}$$

则对每个 k 都有 $p(\lambda_k)=0$, 于是由等式 (9.21) 得

$$p(\boldsymbol{A}) = \boldsymbol{O}. \tag{9.25}$$

这当然就是在 6.13 节中讨论的 Cayley-Hamilton 定理. 上面的论述给出了当矩阵的特征值互不相同时 Cayley-Hamilton 定理的又一个证明.

在展开式 (9.21) 中, 选择一个与上面稍有不同的多项式, 我们可以用它给出 $L_k(A)$ 的显式表示. 我们不用特征多项式 (9.24), 而是选择从特征多项式中去掉一个因子而得到的 $n-1$ 次多项式来代替. 例如, 令 $p_1(\lambda)$ 为从特征多项式 (9.24) 中舍去第一个因子 $(\lambda-\lambda_1)$ 所得到的多项式, 即

$$p_1(\lambda)=(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

于是当 $k\geqslant 2$ 时都有 $p_1(\lambda_k)=0$, 而

$$p_1(\lambda_1)=(\lambda_1-\lambda_2)\cdots(\lambda_1-\lambda_n).$$

注意到由于各 λ_k 互不相同, 因此 $p_1(\lambda_1)\neq 0$. 在展开式 (9.21) 中使用这个多项式, 我们得

$$p_1(\boldsymbol{A})=\sum_{k=1}^{n}p_1(\lambda_k)L_k(\boldsymbol{A})=p_1(\lambda_1)L_1(\boldsymbol{A}).$$

由于 $p_1(\lambda_1)\neq 0$, 因此可以从上式求解 $L_1(A)$ 得

$$L_1(\boldsymbol{A})=\frac{p_1(\boldsymbol{A})}{p_1(\lambda_1)}=\prod_{j=2}^{n}\frac{\boldsymbol{A}-\lambda_j\boldsymbol{I}}{\lambda_1-\lambda_j}.$$

类似地, 如果令 $p_2(\lambda)$ 表示从特征多项式 (9.24) 中舍去第二个因子 $(\lambda-\lambda_2)$ 而得到的多项式, 即若

$$p_2(\lambda)=\prod_{\substack{j=1\\ j\neq 2}}^{n}(\lambda-\lambda_j),$$

我们看到当 $k\neq 2$ 时都有 $p_2(\lambda_k)=0$, 而

$$p_2(\lambda_2)=\prod_{\substack{j=1\\ j\neq 2}}^{n}(\lambda_2-\lambda_j)\neq 0.$$

将多项式 $p_2(\lambda)$ 用于展开式 (9.21), 我们得

$$p_2(\boldsymbol{A})=\sum_{k=1}^{n}p_2(\lambda_k)L_k(\boldsymbol{A})=p_2(\lambda_2)L_2(\boldsymbol{A}).$$

从上式求 $L_2(\boldsymbol{A})$ 得

$$L_2(\boldsymbol{A})=\frac{p_2(\boldsymbol{A})}{p_2(\lambda_2)}=\prod_{\substack{j=1\\ j\neq 2}}^{n}\frac{\boldsymbol{A}-\lambda_j\boldsymbol{I}}{\lambda_2-\lambda_j}.$$

同理可得, 对任意 k, 我们都有关于 $L_k(\boldsymbol{A})$ 的一般公式 (9.23). □

例 1 若 $\boldsymbol{A}$ 为 3×3 矩阵, $\boldsymbol{A}$ 有 3 个不同的特征值 λ,μ,ν, 则由定理 9.9 得

$$e^{t\boldsymbol{A}}=e^{\lambda t}\frac{(\boldsymbol{A}-\mu\boldsymbol{I})(\boldsymbol{A}-\nu\boldsymbol{I})}{(\lambda-\mu)(\lambda-\nu)}+e^{\mu t}\frac{(\boldsymbol{A}-\lambda\boldsymbol{I})(\boldsymbol{A}-\nu\boldsymbol{I})}{(\mu-\lambda)(\mu-\nu)}+e^{\nu t}\frac{(\boldsymbol{A}-\lambda\boldsymbol{I})(\boldsymbol{A}-\mu\boldsymbol{I})}{(\nu-\lambda)(\nu-\mu)}.$$ □

例 2 对 2×2 矩阵 $\boldsymbol{A}=\begin{bmatrix}5&4\\1&2\end{bmatrix}$ 求 $e^{t\boldsymbol{A}}$.

解 矩阵 $\boldsymbol{A}$ 有两个不同的特征值 $\lambda_1=6,\lambda_2=1$, 因此由定理 9.9 得

$$e^{tA} = e^{6t}L_1(A) + e^{t}L_2(A),$$

其中

$$L_1(A) = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2} = \frac{1}{5}(A - I) = \frac{1}{5}\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix},$$

$$L_2(A) = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1} = \frac{1}{-5}(A - 6I) = \frac{1}{-5}\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix}.$$

从而得

$$e^{tA} = \frac{e^{6t}}{5}\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} + \frac{e^{t}}{-5}\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4e^{6t} + e^{t} & 4e^{6t} - 4e^{t} \\ e^{6t} - e^{t} & e^{6t} + 4e^{t} \end{bmatrix}.$$ □

例 3　求线性齐次微分方程组

$$\begin{cases} y_1' = 5y_1 + 4y_2 \\ y_2' = y_1 + 2y_2 \end{cases}$$

满足初始条件 $y_1(0) = 2$ 与 $y_2(0) = 3$ 的解.

解法 1　原方程组可表示成如下矩阵形式:

$$Y'(t) = AY(t),\ Y(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \text{此处 } A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}.$$

由定理 9.7 可得其解为 $Y(t) = e^{tA}Y(0)$. 利用例 2 中求出的矩阵 e^{tA}, 我们得

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4e^{6t} + e^{t} & 4e^{6t} - 4e^{t} \\ e^{6t} - e^{t} & e^{6t} + 4e^{t} \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

由上式即得

$$y_1 = 4e^{6t} - 2e^{t}, \quad y_2 = e^{6t} + 2e^{t}. \tag{9.26}$$

解法 2　还有一种不需求出 e^{tA} 的显式表示的解法. 由于 A 可对角化, 所以 $e^{tA} = Ce^{tD}C^{-1}$, 此处

$$e^{tD} = \begin{bmatrix} e^{6t} & 0 \\ 0 & e^{t} \end{bmatrix}.$$

于是无需实际计算 e^{tA}, 就可以断定它的各个元素都是 e^{6t} 与 e^{t} 的常系数线性组合. 又由于 $Y(t) = e^{tA}Y(0)$, 所以解向量 $Y(t)$ 的各个元素也都是 e^{6t} 与 e^{t} 的常系数线性组合. 因此这些解本身必可表为如下形式:

$$y_1 = ae^{6t} + be^{t}, \quad y_2 = ce^{6t} + de^{t},$$

其中 a, b, c, d 为待定的常数系数. 我们可以利用例题中给出的两个微分方程及两个初始条件来确定这 4 个待定常数. 事实上, 由初始条件可得 $a + b = 2, c + d = 3$, 而由微分方程可得 $a = 4c, b = -d$. 于是得 $a = 4, b = -2, c = 1, d = 2$. 和前面得到的解 (9.26) 相同. □

注　例 3 的解法 2 表明, 当矩阵 A 可以对角化时, 我们可以采用另一种解法来求解齐次线性微分方程组.

$$y'(t) = Ay(t), y(0) = B.$$

如果 A 的特征值为 $\lambda_1, \cdots, \lambda_n$, 则指数矩阵 e^{tA} 的各个元素都是 $e^{t\lambda_1}, \cdots, e^{t\lambda_n}$ 的常系数线性组合, 从而解矩阵 $Y(t)$ 的各元素也都是这样的线性组合, 因此我们可以将作为解 $Y(t)$ 的各元素的函数 $y_{ij}(t)$ 表示为 $e^{t\lambda_1}, \cdots, e^{t\lambda_n}$ 的待定常数系数的线性组合, 然后利用微分方程 $Y'(t) = AY(t)$ 及初始条件 $Y(0) = B$ 来确定这些待定系数.

当 A 不能对角化时, 已经有多种方法可以用来计算 e^{tA}, 大多数方法要求作预备变换, 其本质决定于 A 的特征值的重数. 在 9.12 节中, 我们介绍计算 e^{tA} 的一种实用的直接方法, 不论 A 是否可对角化, 这种方法对任何方阵 A 都能适用, 它不需要作任何种类的预备变换. 这个方法由 E. J. Putzer 在《美国数学月刊》, Vol.73(1966), 99 2-7 中提出. 这个方法的基础是 Cayley-Hamilton 定理.

9.11 习 题

在习题 1 ~ 4 中, (a) 将 A^{-1}, A^2 及所有 A 的更高次幂表示成 I 与 A 的线性组合 (可以应用 Cayley-Hamilton 定理); (b) 求出 e^{tA}.

1. $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.　　2. $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$.　　3. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.　　4. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

在习题 5 ~ 7 中, (a) 计算 A^n, 将 A^3 表示成 I, A 与 A^2 的线性组合; (b) 求出 e^{tA}.

5. $A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.　　6. $A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.　　7. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

8. (a) 若 $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. 证明: $e^{tA} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$.

   (b) 当 $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ 且 a 与 b 为实数时对 e^{tA} 给出相应的表达式.

9. 设 $F(t) = \begin{bmatrix} t & t-1 \\ 0 & 1 \end{bmatrix}$. 证明: $e^{F(t)} = eF(e^{t-1})$.

10. 设 $A(t)$ 是 t 的纯量函数, $e^{A(t)}$ 的导数是 $e^{A(t)}A'(t)$. 计算当 $A(t) = \begin{bmatrix} 1 & t \\ 0 & 0 \end{bmatrix}$ 时 $e^{A(t)}$ 的导数并证明它既不等于乘积 $e^{A(t)}A'(t)$ 也不等于乘积 $A'(t)e^{A(t)}$.

11. 设 $A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, 将 e^{tA} 表示成 I, A, A^2 的线性组合.

12. 设 $A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, 证明: $e^{A} = \begin{bmatrix} x^2 & xy & y^2 \\ 2xy & x^2+y^2 & 2xy \\ y^2 & xy & x^2 \end{bmatrix}$, 此处 $x = \cosh 1, y = \sinh 1$. 其中 $\cosh t = \frac{1}{2}(e^t + e^{-t})$ 与 $\sinh t = \frac{1}{2}(e^t - e^{-t})$ 为双曲函数.

13. 举例指出等式 $e^{A+B} = e^A e^B$ 并不是对任意矩阵指数都成立的. 设 $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ 与 $B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$. 计算矩阵 $e^A e^B, e^B e^A, e^{A+B}$, 验证这 3 个结果各不相同.

14. 给定 $n\times n$ 常数矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$.
    (a) 若 $e^{t\boldsymbol{A}}e^{t\boldsymbol{B}}=e^{t\boldsymbol{B}}e^{t\boldsymbol{A}}$ 对所有实数 t 都成立, 证明: $\boldsymbol{AB}=\boldsymbol{BA}$.(建议: 求导)
    (b) 今设对所有实数 t 都有 $e^{t(\boldsymbol{A}+\boldsymbol{B})}=e^{t\boldsymbol{A}}e^{t\boldsymbol{B}}$, 证明: $\boldsymbol{AB}=\boldsymbol{BA}$, 或举出反例说明该等式不成立.
15. 给定 $n\times n$ 矩阵 (a_{ij}), 其中当 $i+j=n+1$ 时 $a_{ij}=n$, 而在其他情形都有 $a_{ij}=0$.
    (a) 证明: 对每一个 $k\geqslant 1$ 都有 $\boldsymbol{A}^{2k}=n^{2k}\boldsymbol{A}$ 及 $\boldsymbol{A}^{2k+1}=n^{2k}\boldsymbol{A}$.
    (b) 用 $n,\boldsymbol{A}$ 与 $\boldsymbol{I}$ 表示 $e^{t\boldsymbol{A}}$, 当 $n=2$ 时应有

$$e^{t\boldsymbol{A}}=(\cosh 2t)\boldsymbol{I}+\frac{1}{2}(\sinh 2t)\boldsymbol{A}.$$

16. 给定 $n\times n$ 矩阵 $\boldsymbol{A}$ 使对某些选定的纯量 a 与 b 有 $\boldsymbol{A}^2=a\boldsymbol{I}+b\boldsymbol{A}$. 则 $\boldsymbol{A}$ 的高于 2 次的幂也是 $\boldsymbol{I}$ 与 $\boldsymbol{A}$ 的线性组合, 从而指数矩阵 $e^{t\boldsymbol{A}}$ 也可简化为形式

$$e^{t\boldsymbol{A}}=f(t)\boldsymbol{I}+g(t)\boldsymbol{A},$$

此处 $f(t)$ 与 $g(t)$ 都是 t 的纯量函数.
    (a) 设 $\boldsymbol{I}$ 与 $\boldsymbol{A}$ 线性无关, 证明: $f(t)$ 与 $g(t)$ 都满足二阶微分方程

$$y''=ay+by'.$$

    (b) 在 $\boldsymbol{A}^2=4\boldsymbol{I}+3\boldsymbol{A}$ 的特殊情形, 给出 $f(t)$ 与 $g(t)$ 的显式表示, 使得

$$e^{tA}=f(t)I+g(t)A.$$

9.12　计算 e^{tA} 的 Putzer 方法

Cayley-Hamilton 定理表明, 任一 $n\times n$ 矩阵 $\boldsymbol{A}$ 的 n 次幂都可以表示成 $\boldsymbol{I},\boldsymbol{A},\boldsymbol{A}^2,\cdots,\boldsymbol{A}^{n-1}$ 的线性组合. 因此任一更高次幂 $\boldsymbol{A}^{n+1}$, $\boldsymbol{A}^{n+2}$, $\cdots$ 也都可以表示为 $\boldsymbol{I}$, $\boldsymbol{A}$, $\boldsymbol{A}^2$, $\cdots$, $\boldsymbol{A}^{n-1}$ 的线性组合. 从而在定义 $e^{t\boldsymbol{A}}$ 的无穷级数中, 当 $k\geqslant n$ 时, 各项 $\frac{t^k\boldsymbol{A}^k}{k!}$ 也都是 $t^k\boldsymbol{I},t^k\boldsymbol{A},t^k\boldsymbol{A}^2,\cdots,t^k\boldsymbol{A}^{n-1}$ 的线性组合. 由此可知 $e^{t\boldsymbol{A}}$ 可以表示成 $\boldsymbol{A}$ 的形如

$$e^{t\boldsymbol{A}}=\sum_{k=0}^{n-1}q_k(t)\boldsymbol{A}^k \tag{9.27}$$

的多项式, 其中纯量系数 $q_k(t)$ 依赖于 t. E. J. Putzer 给出了将 $e^{t\boldsymbol{A}}$ 表示为 $\boldsymbol{A}$ 的多项式的两个有用的方法, 下述定理给出其中较为简单的那个方法.

定理 9.10　设 $\lambda_1,\cdots,\lambda_n$ 为 $n\times n$ 矩阵 $\boldsymbol{A}$ 的特征值, 定义 $\boldsymbol{A}$ 的多项式序列如下:

$$P_0(\boldsymbol{A})=\boldsymbol{I},P_k(\boldsymbol{A})=\prod_{m=1}^{k}(\boldsymbol{A}-\lambda_m\boldsymbol{I}),\ 对\ k=1,2,\cdots,n. \tag{9.28}$$

则我们有

$$e^{t\boldsymbol{A}}=\sum_{k=0}^{n-1}r_{k+1}(t)P_k(\boldsymbol{A}), \tag{9.29}$$

此处纯量系数 $r_1(t),\cdots,r_n(t)$ 由下述三角形一阶线性微分方程组递归地给出:

$$\begin{aligned}&r_1'(t)=\lambda_1r_1(t),\qquad r_1(0)=1,\\&r_{k+1}'(t)=\lambda_{k+1}r_{k+1}(t)+r_k(t),\quad r_{k+1}(0)=0,\quad (k=1,2,\cdots,n-1).\end{aligned} \tag{9.30}$$

注 展开式 (9.29) 并没有像展开式 (9.27) 中那样直接将 e^{tA} 按 A 的幂展开, 而是将 e^{tA} 表示成多项式 $P_0(A), P_1(A), \cdots, P_{n-1}(A)$ 的一个线性组合. 只要知道 A 的特征值, 这些多项式就容易求得. 由于微分方程组 (9.30) 是三角形的, 因此 (9.29) 中的乘子 $r_1(t), \cdots, r_n(t)$ 也很容易确定. 当这些解逐个确定后, 每个微分方程只包含一个未知函数.

证明 设 $r_1(t), \cdots, r_n(t)$ 是由方程组 (9.30) 确定的纯量函数, 再由

$$F(t) = \sum_{k=0}^{n-1} r_{k+1}(t) P_k(A) \tag{9.31}$$

定义一个矩阵函数 F. 我们将证明 F 和 e^{tA} 满足同一个微分方程, 即 $F'(t) = AF(t)$ 以及在 $t=0$ 处的同一个初始条件, 由此即可得 $F(t) = e^{tA}$.

首先, 我们指出 $F(0) = r_1(0)P_0(A) = I$. 其次, 对等式 (9.31) 两边微分并利用方程组 (9.30) 中的递推公式得

$$F'(t) = \sum_{k=0}^{n-1} r'_{k+1}(t) P_k(A) = \sum_{k=0}^{n-1} \{r_k(t) + \lambda_{k+1} r_{k+1}(t)\} P_k(A),$$

此处 $r_0(t)$ 定义为 0. 将上式重写为

$$F'(t) = \sum_{k=0}^{n-2} r_{k+1}(t) P_{k+1}(A) + \sum_{k=0}^{n-1} \lambda_{k+1} r_{k+1}(t) P_k(A),$$

然后减去 $\lambda_n F(t) = \sum\limits_{k=0}^{n-1} \lambda_n r_{k+1}(t) P_k(A)$ 即得如下关系式:

$$F'(t) - \lambda_n F(t) = \sum_{k=0}^{n-2} r_{k+1}(t) \{P_{k+1}(A) + (\lambda_{k+1} - \lambda_n) P_k(A)\}. \tag{9.32}$$

然后由 (9.28) 得 $P_{k+1}(A) = (A - \lambda_{k+1} I) P_k(A)$, 因此

$$\begin{aligned} P_{k+1}(A) + (\lambda_{k+1} - \lambda_n) P_k(A) &= (A - \lambda_{k+1} I) P_k(A) + (\lambda_{k+1} - \lambda_n) P_k(A) \\ &= (A - \lambda_n I) P_k(A). \end{aligned}$$

于是等式 (9.32) 变成

$$\begin{aligned} F'(t) - \lambda_n F(t) &= (A - \lambda_n I) \sum_{k=0}^{n-2} r_{k+1}(t) P_k(A) \\ &= (A - \lambda_n I)\{F(t) - r_n(t) P_{n-1}(A)\} \\ &= (A - \lambda_n I) F(t) - r_n(t) P_n(A). \end{aligned}$$

由 Cayley-Hamilton 定理得 $P_n(A) = O$, 因而上式变成

$$F'(t) - \lambda_n F(t) = (A - \lambda_n I) F(t) = AF(t) - \lambda_n F(t),$$

由此即可得 $F'(t) = AF(t)$. 由于 $F(0) = I$, 因此由唯一性定理即得 $F(t) = e^{tA}$. □

例 1 设 A 为 2×2 矩阵, 它的两个根都等于 λ, 将 e^{tA} 表示成 I 与 A 的线性组合.

解 令 $\lambda_1 = \lambda_2 = \lambda$, 考虑一阶微分方程组

$$\begin{cases} r_1'(t) = \lambda r_1(t), & r_1(0) = 1, \\ r_2'(t) = \lambda r_2(t) + r_1(t), & r_2(0) = 0. \end{cases}$$

顺次解上面这两个微分方程, 我们得

$$r_1(t) = \mathrm{e}^{\lambda t}, \qquad r_2(t) = t\mathrm{e}^{\lambda t}.$$

由于 $P_0(\boldsymbol{A}) = \boldsymbol{I}$ 及 $P_1(\boldsymbol{A}) = \boldsymbol{A} - \lambda\boldsymbol{I}$, 因此要求的关于 $\mathrm{e}^{t\boldsymbol{A}}$ 的公式是

$$\mathrm{e}^{t\boldsymbol{A}} = \mathrm{e}^{\lambda t}\boldsymbol{I} + t\mathrm{e}^{\lambda t}(\boldsymbol{A} - \lambda\boldsymbol{I}) = \mathrm{e}^{\lambda t}(1 - \lambda t)\boldsymbol{I} + t\mathrm{e}^{\lambda t}\boldsymbol{A}. \qquad \square \quad (9.33)$$

例 2 在例 1 中设 $\boldsymbol{A}$ 的特征值为 λ 与 μ, $\lambda \neq \mu$, 将 $\mathrm{e}^{t\boldsymbol{A}}$ 表示成 $\boldsymbol{I}$ 与 $\boldsymbol{A}$ 的线性组合.

解 在这一情形, 微分方程组为

$$\begin{cases} r_1'(t) = \lambda r_1(t), & r_1(0) = 1, \\ r_2'(t) = \mu r_2(t) + r_1(t), & r_2(0) = 0. \end{cases}$$

它的解由

$$r_1(t) = \mathrm{e}^{\lambda t}, \qquad r_2(t) = \frac{\mathrm{e}^{\lambda t} - \mathrm{e}^{\mu t}}{\lambda - \mu}$$

给出. 由于 $P_0(\boldsymbol{A}) = \boldsymbol{I}$ 及 $P_1(\boldsymbol{A}) = \boldsymbol{A} - \lambda\boldsymbol{I}$, 因此 $\mathrm{e}^{t\boldsymbol{A}}$ 的公式为

$$\mathrm{e}^{t\boldsymbol{A}} = \mathrm{e}^{\lambda t}\boldsymbol{I} + \frac{\mathrm{e}^{\lambda t} - \mathrm{e}^{\mu t}}{\lambda - \mu}(\boldsymbol{A} - \lambda\boldsymbol{I}) = \frac{\lambda\mathrm{e}^{\mu t} - \mu\mathrm{e}^{\lambda t}}{\lambda - \mu}\boldsymbol{I} + \frac{\mathrm{e}^{\lambda t} - \mathrm{e}^{\mu t}}{\lambda - \mu}\boldsymbol{A}. \qquad (9.34)$$

读者可自行验证这个解与定理 9.9 中当 $n = 2$ 时给出的结果是一致的. $\square$

应当指出, 关于 $\mathrm{e}^{t\boldsymbol{A}}$ 的公式 (9.33) 可以看作公式 (9.34) 中令 $\mu \to \lambda$ 的极限情形 (见 9.14 节中的习题 8).

若特征值 λ 与 μ 为共轭复数, 则 $\mathrm{e}^{\lambda t}$ 与 $\mathrm{e}^{\mu t}$ 也是复数. 但当 λ 与 μ 为共轭复数时, 在公式 (9.34) 中 $\boldsymbol{I}$ 与 $\boldsymbol{A}$ 的纯量倍将全是实的. 例如, 假定

$$\lambda = \alpha + \mathrm{i}\beta, \quad \mu = \alpha - \mathrm{i}\beta, \quad 此处\beta \neq 0,$$

则 $\lambda - \mu = 2i\beta$, 因此由公式 (9.34) 得

$$\begin{aligned} \mathrm{e}^{t\boldsymbol{A}} &= \mathrm{e}^{(\alpha+\mathrm{i}\beta)t}\boldsymbol{I} + \frac{\mathrm{e}^{(\alpha+\mathrm{i}\beta)t} - \mathrm{e}^{(\alpha-i\beta)t}}{2i\beta}\{\boldsymbol{A} - (\alpha + \mathrm{i}\beta)\boldsymbol{I}\} \\ &= \mathrm{e}^{\alpha t}\left\{\mathrm{e}^{\mathrm{i}\beta t}\boldsymbol{I} + \frac{\mathrm{e}^{\mathrm{i}\beta t} - \mathrm{e}^{-\mathrm{i}\beta t}}{2\mathrm{i}\beta}(\boldsymbol{A} - \alpha\boldsymbol{I} - \mathrm{i}\beta\boldsymbol{I})\right\} \\ &= \mathrm{e}^{\alpha t}\left\{(\cos\beta t + \mathrm{i}\sin\beta t)\boldsymbol{I} + \frac{\sin\beta t}{\beta}(\boldsymbol{A} - \alpha\boldsymbol{I} - \mathrm{i}\beta\boldsymbol{I})\right\}. \end{aligned}$$

由于其中含有 i 的两项正好抵消, 因此得

$$\mathrm{e}^{t\boldsymbol{A}} = \frac{\mathrm{e}^{\alpha t}}{\beta}\{(\beta\cos\beta t - \alpha\sin\beta t)\boldsymbol{I} + \sin\beta t\boldsymbol{A}\}. \qquad (9.35)$$

例 3 设 $\boldsymbol{A}$ 为 4×4 矩阵, $\boldsymbol{A}$ 的特征值为 $\lambda_1 = \lambda_2 = 0, \lambda_3 = \lambda_4 = 1$. 将 $\mathrm{e}^{t\boldsymbol{A}}$ 表示成 $\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2$ 与 $\boldsymbol{A}^3$ 的线性组合.

解 解三角形微分方程组 (9.30) 得

$$r_1(t)=1,\ r_2(t)=t,\ r_3(t)=e^t-t-1,\ r_4(t)=(t-2)e^t+t+2.$$

由定理 9.10 则得

$$\begin{aligned}e^{tA}&=I+tA+(e^t-t-1)A^2+\{(t-2)e^t+t+2\}A^2(A-I)\\&=I+tA+\{(3-t)e^t-2t-3\}A^2+\{(t-2)e^t+t+2\}A^3.\end{aligned}$$ □

9.13 在特殊情形下计算 e^{tA} 的方法

将 e^{tA} 表示为 A 的多项式的 Putzer 方法是完全一般的方法, 因为它对所有的方阵 A 都适用. 例如定理 9.9 提供了当 A 的所有特征值都不相同时计算 e^{tA} 的一个直接方法. 这一节将给出在两种特殊情形下计算 e^{tA} 的比较简便的方法: (a)A 的所有特征值都相等; (b)A 有两个不同的特征值, 而且其中一个的重数是 1.

定理 9.11 设 A 为 $n\times n$ 矩阵, A 的特征值都等于 λ, 则我们有

$$e^{tA}=e^{\lambda t}\sum_{k=0}^{n-1}\frac{t^k}{k!}(A-\lambda I)^k. \tag{9.36}$$

证明 由于矩阵 λtI 与 $t(A-\lambda I)$ 可交换, 因此

$$e^{tA}=e^{\lambda tI}e^{t(A-\lambda I)}=\left(e^{\lambda t}I\right)\sum_{k=0}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k.$$

由 Cayley-Hamilton 定理得, 对任意 $k\geqslant n$ 都有 $(A-\lambda I)^k=O$, 因此上式右边的级数到 $k=n-1$ 时终止, 即得结论. □

定理 9.12 设 $n\geqslant 3$, A 为 $n\times n$ 矩阵, A 有两个不同的特征值 λ 与 μ, 其中 λ 的重数为 $n-1$, μ 的重数为 1, 则我们有

$$e^{tA}=e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k+\left\{\frac{e^{\mu t}}{(\mu-\lambda)^{n-1}}-\frac{e^{\lambda t}}{(\mu-\lambda)^{n-1}}\sum_{k=0}^{n-2}\frac{t^k}{k!}(\mu-\lambda)^k\right\}(A-\lambda I)^{n-1}.$$

证明 如同定理 9.11 的证明一样, 我们先把 e^{tA} 改写成如下形式:

$$\begin{aligned}e^{tA}&=e^{\lambda t}\sum_{k=0}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k=e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k+e^{\lambda t}\sum_{k=n-1}^{\infty}\frac{t^k}{k!}(A-\lambda I)^k\\&=e^{\lambda t}\sum_{k=0}^{n-2}\frac{t^k}{k!}(A-\lambda I)^k+e^{\lambda t}\sum_{r=0}^{\infty}\frac{t^{n-1+r}}{(n-1+r)!}(A-\lambda I)^{n-1+r}.\end{aligned}$$

然后应用 Cayley-Hamilton 定理来精确估计上述级数中实际出现的项数. 由于

$$A-\mu I=A-\lambda I-(\mu-\lambda)I,$$

因此有

$$(A-\lambda I)^{n-1}(A-\mu I)=(A-\lambda I)^n-(\mu-\lambda)(A-\lambda I)^{n-1}.$$

由 Cayley-Hamilton 定理, 上式左边等于 O, 因此

$$(\boldsymbol{A}-\lambda \boldsymbol{I})^{n}=(\mu-\lambda)(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1}.$$

重复利用这个关系 r 次可得

$$(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1+r}=(\mu-\lambda)^{r}(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1}.$$

于是得级数

$$\begin{aligned}
&\sum_{r=0}^{\infty} \frac{t^{n-1+r}}{(n-1+r)!}(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1+r} \\
=&\frac{1}{(\mu-\lambda)^{n-1}} \sum_{k=n-1}^{\infty} \frac{t^{k}}{k!}(\mu-\lambda)^{k}(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1} \\
=&\frac{1}{(\mu-\lambda)^{n-1}}\left\{\mathrm{e}^{t(\mu-\lambda)}-\sum_{k=0}^{n-2} \frac{t^{k}}{k!}(\mu-\lambda)^{k}\right\}(\boldsymbol{A}-\lambda \boldsymbol{I})^{n-1}.
\end{aligned}$$

这就完成了定理的证明. 定理 9.12 也可用 Putzer 方法得到, 不过推导过程要复杂得多. □

3 阶矩阵 定理 9.11、定理 9.12 与定理 9.9 中的显式公式覆盖了阶数 $n \leqslant 3$ 的所有矩阵. 由于 3×3 矩阵在实际中常常出现, 因此为方便参阅, 我们把 $n=3$ 时的公式罗列如下.

情形 1 $\boldsymbol{A}$ 为 3×3 矩阵, $\boldsymbol{A}$ 的 3 个根为 $\lambda, \lambda, \lambda$, 则由定理 9.11 有

$$\mathrm{e}^{t \boldsymbol{A}}=\mathrm{e}^{\lambda t}\left\{\boldsymbol{I}+t(\boldsymbol{A}-\lambda \boldsymbol{I})+\frac{1}{2} t^{2}(\boldsymbol{A}-\lambda \boldsymbol{I})^{2}\right\}.$$

情形 2 $\boldsymbol{A}$ 为 3×3 矩阵, $\boldsymbol{A}$ 的特征值为 λ, λ, μ, 此处 $\lambda \neq \mu$, 则由定理 9.12 得

$$\mathrm{e}^{t \boldsymbol{A}}=\mathrm{e}^{\lambda t}\left\{\boldsymbol{I}+t(\boldsymbol{A}-\lambda \boldsymbol{I})+\frac{\mathrm{e}^{\mu t}-\mathrm{e}^{\lambda t}}{(\mu-\lambda)^{2}}(\boldsymbol{A}-\lambda \boldsymbol{I})^{2}-\frac{t \mathrm{e}^{\lambda t}}{\mu-\lambda}(\boldsymbol{A}-\lambda \boldsymbol{I})^{2}\right\}.$$

情形 3 $\boldsymbol{A}$ 为 3×3 矩阵, $\boldsymbol{A}$ 有 3 个不同的特征值 λ, μ, ν, 则由定理 9.9 得

$$\mathrm{e}^{t \boldsymbol{A}}=\mathrm{e}^{\lambda t} \frac{(\boldsymbol{A}-\mu \boldsymbol{I})(\boldsymbol{A}-\nu \boldsymbol{I})}{(\lambda-\mu)(\lambda-\nu)}+\mathrm{e}^{\mu t} \frac{(\boldsymbol{A}-\lambda \boldsymbol{I})(\boldsymbol{A}-\nu \boldsymbol{I})}{(\mu-\lambda)(\mu-\nu)}+\mathrm{e}^{\nu t} \frac{(\boldsymbol{A}-\lambda \boldsymbol{I})(\boldsymbol{A}-\mu \boldsymbol{I})}{(\nu-\lambda)(\nu-\mu)}.$$

例 设 $\boldsymbol{A}=\begin{bmatrix}0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4\end{bmatrix}$, 求 $\mathrm{e}^{t \boldsymbol{A}}$.

解 $\boldsymbol{A}$ 的特征值为 $1,1,2$, 因此由情形 2 的公式得

$$\mathrm{e}^{t \boldsymbol{A}}=\mathrm{e}^{t}\{\boldsymbol{I}+t(\boldsymbol{A}-\boldsymbol{I})\}+\left(\mathrm{e}^{2 t}-\mathrm{e}^{t}\right)(\boldsymbol{A}-\boldsymbol{I})^{2}-t \mathrm{e}^{t}(\boldsymbol{A}-\boldsymbol{I})^{2}. \tag{9.37}$$

将 $\boldsymbol{A}$ 的同次幂整理在一起, 得

$$\mathrm{e}^{t \boldsymbol{A}}=\left(-2 t \mathrm{e}^{t}+\mathrm{e}^{2 t}\right) \boldsymbol{I}+\left\{(3 t+2) \mathrm{e}^{t}-2 \mathrm{e}^{2 t}\right\} \boldsymbol{A}-\left\{(t+1) \mathrm{e}^{t}-\mathrm{e}^{2 t}\right\} \boldsymbol{A}^{2}. \tag{9.38}$$

现在, 我们计算 $(\boldsymbol{A}-\boldsymbol{I})^{2}$ 或 $\boldsymbol{A}^{2}$. 然后利用等式 (9.37) 或 (9.38) 即可将 $\mathrm{e}^{t \boldsymbol{A}}$ 表示为如下 3×3 矩阵

$$e^{tA}=\begin{bmatrix}-2te^t+e^{2t} & (3t+2)e^t-2e^{2t} & -(t+1)e^t+e^{2t}\\ -2(t+1)e^t+2e^{2t} & (3t+5)e^t-4e^{2t} & -(t+2)e^t+2e^{2t}\\ -2(t+2)e^t+4e^{2t} & (3t+8)e^t-8e^{2t} & -(t+4)e^t+4e^{2t}\end{bmatrix}. \qquad \square$$

9.14 习　　题

对习题 1 ~ 6 中的矩阵, 将 e^{tA} 表示为 A 的多项式.

1. $A=\begin{bmatrix}5 & -2\\ 4 & -1\end{bmatrix}$.

2. $A=\begin{bmatrix}1 & 2\\ 2 & -1\end{bmatrix}$.

3. $A=\begin{bmatrix}1 & 0 & 2\\ 0 & 1 & 3\\ 0 & 0 & 1\end{bmatrix}$.

4. $A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -6 & -11 & -6\end{bmatrix}$.

5. $A=\begin{bmatrix}3 & -1 & 1\\ 2 & 0 & 1\\ 1 & -1 & 2\end{bmatrix}$.

6. $A=\begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 2 & 1 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 4\end{bmatrix}$.

7. (a) 已知 3×3 矩阵 A 的所有特征值都等于 λ. 证明:

$$e^{tA}=\frac{1}{2}e^{\lambda t}\left\{\left(\lambda^2t^2-2\lambda t+2\right)I+\left(-2\lambda t^2+2t\right)A+t^2A^2\right\}.$$

(b) 若 A 为所有特征值都等于 λ 的 4×4 矩阵, 给出相应的 e^{tA} 的公式.

8. 证明关于 e^{tA} 的公式 (9.33) 可以看作从公式 (9.34) 取 $\mu\to\lambda$ 时的极限情形. [提示: 对 0/0 不定型应用 L'Hôspital 法则.]

在习题 9 ~ 16 中, 求微分方程组 $Y'=AY$ 对给定初始条件的解

9. $A=\begin{bmatrix}1 & 2\\ 2 & -1\end{bmatrix}, Y(0)=\begin{bmatrix}c_1\\ c_2\end{bmatrix}$.

10. $A=\begin{bmatrix}-5 & 3\\ -15 & 7\end{bmatrix}, Y(0)=\begin{bmatrix}1\\ 1\end{bmatrix}$.

11. $A=\begin{bmatrix}3 & -1 & 1\\ 2 & 0 & 1\\ 1 & -1 & 2\end{bmatrix}, Y(0)=\begin{bmatrix}1\\ -1\\ 2\end{bmatrix}$.

12. $A=\begin{bmatrix}2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 1\end{bmatrix}, Y(0)=\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix}$.

13. $A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -6 & -11 & -6\end{bmatrix}, Y(0)=\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$.

14. $A=\begin{bmatrix}-2 & 2 & -3\\ 2 & 1 & -6\\ -1 & -2 & 0\end{bmatrix}, Y(0)\begin{bmatrix}8\\ 0\\ 0\end{bmatrix}$.

15. $A=\begin{bmatrix}1 & 1 & 0 & 0\\ 0 & 2 & 1 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 4\end{bmatrix}, Y(0)=\begin{bmatrix}0\\ 0\\ 1\\ 1\end{bmatrix}$.

16. $A=\begin{bmatrix}0 & 0 & 2 & 0\\ 1 & 0 & 0 & 2\\ 0 & 0 & 0 & 4\\ 0 & 0 & 1 & 0\end{bmatrix}, Y(0)\begin{bmatrix}1\\ 0\\ 2\\ 1\end{bmatrix}$.

9.15 常系数非齐次线性微分方程组

下面我们考虑在区间 J 上的非齐次初值问题:

$$Y'(t)=AY(t)+Q(t),\ Y(a)=B. \tag{9.39}$$

此处 $\boldsymbol{A}$ 是 $n\times n$ 常数矩阵, $\boldsymbol{Q}(t)$ 是一个在 J 上连续的 $n\times m$ 矩阵函数, a 是 J 中给定的一点, $\boldsymbol{B}$ 是 $n\times m$ 初值矩阵. 采用和纯量情形相同的程序, 可以得到这个问题的 $n\times m$ 矩阵解的显式表示.

首先, 在方程 (9.39) 两边同乘以指数矩阵 $\mathrm{e}^{-t\boldsymbol{A}}$ 并将微分方程改写成如下形式

$$\mathrm{e}^{-t\boldsymbol{A}}\{\boldsymbol{Y}'(t)-\boldsymbol{A}\boldsymbol{Y}(t)\}=\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Q}(t). \tag{9.40}$$

上式左边是乘积 $\mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Y}(t)$ 的导数. 于是若将等式 (9.40) 两边从 a 到 x 积分, 其中 $x\in J$, 则得

$$\mathrm{e}^{-x\boldsymbol{A}}\boldsymbol{Y}(x)-\mathrm{e}^{-a\boldsymbol{A}}\boldsymbol{Y}(a)=\int_a^x \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Q}(t)\mathrm{d}t.$$

再乘以 $\mathrm{e}^{x\boldsymbol{A}}$ 则得到下述定理中的显式公式 (9.41).

定理 9.13　*设 $\boldsymbol{A}$ 为 $n\times n$ 常数矩阵, $\boldsymbol{Q}$ 为在区间 J 上连续的 $n\times m$ 矩阵函数, 则初值问题*

$$\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t)+\boldsymbol{Q}(t),\ \boldsymbol{Y}(a)=\boldsymbol{B},$$

在 J 上有唯一解, 并且这个解由下述公式给出:

$$\boldsymbol{Y}(x)=\mathrm{e}^{(x-a)\boldsymbol{A}}\boldsymbol{B}+\mathrm{e}^{x\boldsymbol{A}}\int_a^x \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Q}(t)\mathrm{d}t. \qquad \square \tag{9.41}$$

与齐次情形相同, 应用这个公式的困难在于这个指数矩阵的计算.

注意上式右边第一项 $\mathrm{e}^{(x-a)\boldsymbol{A}}\boldsymbol{B}$ 是齐次问题 $\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t)$ 且 $\boldsymbol{Y}(a)=\boldsymbol{B}$ 的解, 第二项是非齐次问题

$$\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t)+\boldsymbol{Q}(t),\ \boldsymbol{Y}(a)=\boldsymbol{O}$$

的解.

我们用例子来说明定理 9.13.

例　在区间 $(-\infty,+\infty)$ 上解初值问题

$$\boldsymbol{Y}'(t)=\boldsymbol{A}\boldsymbol{Y}(t)+\boldsymbol{Q}(t),\ \boldsymbol{Y}(0)=\boldsymbol{B},$$

此处

$$\boldsymbol{A}=\begin{bmatrix}2&-1&1\\0&3&-1\\2&1&3\end{bmatrix},\quad \boldsymbol{Q}(t)=\begin{bmatrix}\mathrm{e}^{2t}\\0\\t\mathrm{e}^{2t}\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

解　根据定理 9.13, 解由下式给出:

$$\boldsymbol{Y}(x)=\mathrm{e}^{x\boldsymbol{A}}\int_0^x \mathrm{e}^{-t\boldsymbol{A}}\boldsymbol{Q}(t)\mathrm{d}t=\int_0^x \mathrm{e}^{(x-t)\boldsymbol{A}}\boldsymbol{Q}(t)\mathrm{d}t. \tag{9.42}$$

$\boldsymbol{A}$ 的特征值是 2, 2 与 4. 为计算 $\mathrm{e}^{x\boldsymbol{A}}$, 我们用 9.12 节情形 2 的公式得

$$\begin{aligned}\mathrm{e}^{x\boldsymbol{A}}&=\mathrm{e}^{2x}\{\boldsymbol{I}+x(\boldsymbol{A}-2\boldsymbol{I})\}+\frac{1}{4}(\mathrm{e}^{4x}-\mathrm{e}^{2x})(\boldsymbol{A}-2\boldsymbol{I})^2-\frac{1}{2}x\mathrm{e}^{2x}(\boldsymbol{A}-2\boldsymbol{I})^2\\&=\mathrm{e}^{2x}\left\{\boldsymbol{I}+x(\boldsymbol{A}-2\boldsymbol{I})+\frac{1}{4}(\mathrm{e}^{2x}-2x-1)(\boldsymbol{A}-2\boldsymbol{I})^2\right\}.\end{aligned}$$

在上式中用 $x-t$ 代替 x 可得 $\mathrm{e}^{(x-t)\boldsymbol{A}}$, 于是积分 (9.42) 中的被积函数为

$$
\begin{aligned}
&\mathrm{e}^{(x-t)\boldsymbol{A}}\boldsymbol{Q}(t)\\
=&\mathrm{e}^{2(x-t)}\left\{\boldsymbol{I}+(x-t)(\boldsymbol{A}-2\boldsymbol{I})+\frac{1}{4}[\mathrm{e}^{2(x-t)}-2(x-t)-1](\boldsymbol{A}-2\boldsymbol{I})^2\right\}\boldsymbol{Q}(t)\\
=&\mathrm{e}^{2x}\begin{bmatrix}1\\0\\t\end{bmatrix}+(\boldsymbol{A}-2\boldsymbol{I})\mathrm{e}^{2x}\begin{bmatrix}x-t\\0\\t(x-t)\end{bmatrix}+\frac{1}{4}(\boldsymbol{A}-2\boldsymbol{I})^2\mathrm{e}^{2x}\begin{bmatrix}\mathrm{e}^{2x}\mathrm{e}^{-2t}-2(x-t)-1\\0\\\mathrm{e}^{2x}t\mathrm{e}^{-2t}-2t(x-t)-t\end{bmatrix}.
\end{aligned}
$$

于是得其积分为

$$
\begin{aligned}
\boldsymbol{Y}(x)&=\int_0^x \mathrm{e}^{(x-t)\boldsymbol{A}}\boldsymbol{Q}(t)\mathrm{d}t\\
&=\mathrm{e}^{2x}\begin{bmatrix}x\\0\\\frac{1}{2}x^2\end{bmatrix}+(\boldsymbol{A}-2\boldsymbol{I})\mathrm{e}^{2x}\begin{bmatrix}\frac{1}{2}x^2\\0\\\frac{1}{6}x^3\end{bmatrix}\\
&\quad+\frac{1}{4}(\boldsymbol{A}-2\boldsymbol{I})^2\mathrm{e}^{2x}\begin{bmatrix}\frac{1}{2}\mathrm{e}^{2x}-\frac{1}{2}-x-x^2\\0\\\frac{1}{4}\mathrm{e}^{2x}-\frac{1}{4}-\frac{1}{2}x-\frac{1}{2}x^2-\frac{1}{3}x^3\end{bmatrix}.
\end{aligned}
$$

由于我们有

$$
\boldsymbol{A}-2\boldsymbol{I}=\begin{bmatrix}0&-1&1\\0&1&-1\\2&1&1\end{bmatrix},\quad (\boldsymbol{A}-2\boldsymbol{I})^2=\begin{bmatrix}2&0&2\\-2&0&-2\\2&0&2\end{bmatrix},
$$

因此得

$$
\begin{aligned}
\boldsymbol{Y}(x)&=\mathrm{e}^{2x}\begin{bmatrix}x\\0\\\frac{1}{2}x^2\end{bmatrix}+\mathrm{e}^{2x}\begin{bmatrix}\frac{1}{6}x^3\\-\frac{1}{6}x^3\\x^2+\frac{1}{6}x^3\end{bmatrix}+\mathrm{e}^{2x}\begin{bmatrix}\frac{3}{8}\mathrm{e}^{2x}-\frac{3}{8}-\frac{3}{4}x-\frac{3}{4}x^2-\frac{1}{6}x^3\\-\frac{3}{8}\mathrm{e}^{2x}+\frac{3}{8}+\frac{3}{4}x+\frac{3}{4}x^2+\frac{1}{6}x^3\\\frac{3}{8}\mathrm{e}^{2x}-\frac{3}{8}-\frac{3}{4}x-\frac{3}{4}x^2-\frac{1}{6}x^3\end{bmatrix}\\
&=\mathrm{e}^{2x}\begin{bmatrix}\frac{3}{8}\mathrm{e}^{2x}-\frac{3}{8}+\frac{1}{4}x-\frac{3}{4}x^2\\-\frac{3}{8}\mathrm{e}^{2x}+\frac{3}{8}+\frac{3}{4}x+\frac{3}{4}x^2\\\frac{3}{8}\mathrm{e}^{2x}-\frac{3}{8}-\frac{3}{4}x+\frac{3}{4}x^2\end{bmatrix}.
\end{aligned}
$$

$\boldsymbol{Y}(x)$ 的各个分量即为所求的函数 y_1, y_2, y_3. □

9.16 习 题

1. 令 $\boldsymbol{Z}$ 为非齐次方程组
$$\boldsymbol{Z}'(t)=\boldsymbol{AZ}(t)+\boldsymbol{Q}(t)$$
在区间 J 上以 $\boldsymbol{Z}(a)$ 为初始值的解. 证明：非齐次方程组
$$\boldsymbol{Y}'(t)=\boldsymbol{AY}(t)+\boldsymbol{Q}(t)$$
在区间 J 上只有一个以 $\boldsymbol{Y}(a)$ 为初始值的解, 并且这个解由下述公式给出：
$$\boldsymbol{Y}(t)=\boldsymbol{Z}(t)+\mathrm{e}^{(t-a)\boldsymbol{A}}\{\boldsymbol{Y}(a)-\boldsymbol{Z}(a)\}.$$

常常可以用一些特殊的方法来确定方程的与给定函数 $\boldsymbol{Q}(t)$ 相类似的特解 $\boldsymbol{Z}(t)$. 习题 2 ~ 3、习题 5 和习题 7 对 $\boldsymbol{Q}(t)=\boldsymbol{C}$, $\boldsymbol{Q}(t)=\mathrm{e}^{\alpha t}\boldsymbol{C}$, $\boldsymbol{Q}(t)=t^m\boldsymbol{C}$ 以及 $\boldsymbol{Q}(t)=(\cos\alpha t)\boldsymbol{C}+(\sin\alpha t)\boldsymbol{D}$ 的情形给出了这样的方法, 此处 $\boldsymbol{C}$ 与 $\boldsymbol{D}$ 是常数向量. 如果这样得到的特解 $\boldsymbol{Z}(t)$ 不具有所要求的初始值, 我们可以如习题 1 所指出的那样对 $\boldsymbol{Z}(t)$ 作修正以得到另一个解 $\boldsymbol{Y}(t)$, 而这个解具有所要求的初始值.

2. (a) 令 $\boldsymbol{A}$ 为一个 $n\times n$ 常数矩阵, $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 为 $\mathbb{R}^n$ 中的常数向量. 证明微分方程组
$$\boldsymbol{Y}'(t)=\boldsymbol{AY}(t)+\boldsymbol{C},\quad \boldsymbol{Y}(a)=\boldsymbol{B}$$
在 $(-\infty,+\infty)$ 中的解由下式给出：
$$Y(x)=\mathrm{e}^{(x-a)\boldsymbol{A}}\boldsymbol{B}+\left(\int_0^{x-a}\mathrm{e}^{u\boldsymbol{A}}\mathrm{d}u\right)\boldsymbol{C}.$$
(b) 若 $\boldsymbol{A}$ 为非奇异矩阵, 证明 (a) 中的积分具有值 $\{\mathrm{e}^{(x-a)\boldsymbol{A}}-\boldsymbol{I}\}\boldsymbol{A}^{-1}$.
(c) 设
$$\boldsymbol{A}=\begin{bmatrix}-1&2\\-2&3\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}1\\2\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}b\\c\end{bmatrix},\quad a=0,$$
求 $\boldsymbol{Y}(x)$ 的显式表示.

3. 给定 $n\times n$ 常数矩阵 $\boldsymbol{A}$, $\mathbb{R}^n$ 中常数向量 $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 以及纯量 α.
(a) 证明非齐次微分方程组 $\boldsymbol{Z}'(t)=\boldsymbol{AZ}(t)+\mathrm{e}^{\alpha t}\boldsymbol{C}$ 存在形如 $\boldsymbol{Z}(t)=\mathrm{e}^{\alpha t}\boldsymbol{B}$ 的解的充要条件是 $(\alpha\boldsymbol{I}-\boldsymbol{A})\boldsymbol{B}=\boldsymbol{C}$.
(b) 若 α 不是 $\boldsymbol{A}$ 的特征值, 证明总可以选取向量 $\boldsymbol{B}$ 使得 (a) 中方程组有形如 $\boldsymbol{Z}(t)=\mathrm{e}^{\alpha t}\boldsymbol{B}$ 的解.
(c) 若 α 不是 $\boldsymbol{A}$ 的特征值, 证明方程组 $\boldsymbol{Y}'(t)=\boldsymbol{AY}(t)+\mathrm{e}^{\alpha t}\boldsymbol{C}$ 的所有解都具有形式 $\boldsymbol{Y}(t)=\mathrm{e}^{t\boldsymbol{A}}(\boldsymbol{Y}(0)-\boldsymbol{B})+\mathrm{e}^{\alpha t}\boldsymbol{B}$, 其中 $\boldsymbol{B}=(\alpha\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{C}$.

4. 用习题 3 中给出的方法求非齐次方程组 $\boldsymbol{Y}'(t)=\boldsymbol{AY}(t)+\mathrm{e}^{2t}\boldsymbol{C}$ 的一组解, 其中
$$\boldsymbol{A}=\begin{bmatrix}3&1\\2&2\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}-1\\-1\end{bmatrix},\quad \boldsymbol{Y}(0)=\begin{bmatrix}0\\1\end{bmatrix}.$$

5. 给定 $n\times n$ 常数矩阵 $\boldsymbol{A}$, $\mathbb{R}^n$ 中常数向量 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 以及正整数 m.
(a) 证明非齐次方程组 $\boldsymbol{Y}'(t)=\boldsymbol{AY}(t)+t^m\boldsymbol{C}$, $\boldsymbol{Y}(0)=\boldsymbol{B}$ 有一组形如
$$\boldsymbol{Y}(t)=\boldsymbol{B}_0+t\boldsymbol{B}_1+t^2\boldsymbol{B}_2+\cdots+t^m\boldsymbol{B}_m,$$
的特解的充要条件为 $\boldsymbol{C}=-\frac{1}{m!}\boldsymbol{A}^{m+1}\boldsymbol{B}$, 其中 $\boldsymbol{B}_0,\boldsymbol{B}_1,\cdots,\boldsymbol{B}_m$ 为常数向量, 并计算此时系数 $\boldsymbol{B}_0,\boldsymbol{B}_1,\cdots,\boldsymbol{B}_m$ 的值.
(b) 若 $\boldsymbol{A}$ 非奇异, 证明总可以找到初始向量 $\boldsymbol{B}$ 使得 (a) 中方程组有一组具有给定形式的解.

6. 考虑非齐次方程组
$$\begin{cases}y_1'=3y_1+y_2+t^2\\ y_2'=2y_1+2y_2+t^3.\end{cases}$$
(a) 求一组具有形式 $\boldsymbol{Y}(t)=\boldsymbol{B}_0+t\boldsymbol{B}_1+t^2\boldsymbol{B}_2+t^3\boldsymbol{B}_3$ 的特解.

(b) 求方程组的一组解使得 $y_1(0) = y_2(0) = 1$.

7. 给定 $n \times n$ 常数矩阵 A, $\mathbb{R}^n$ 中的常数向量 B, C, D 以及非零实数 α, 证明非齐次方程组

$$Y'(t) = AY(t) + (\cos\alpha t)C + (\sin\alpha t)D, \quad Y(0) = B,$$

有一组形如 $Y(t) = (\cos\alpha t)E + (\sin\alpha t)F$ 的特解的充要条件是 $(A^2 + \alpha^2 I)B = -(AC + \alpha D)$, 其中 E 和 F 为常数向量. 用 A, B, C 表示该解中的 E 和 F. 注意到, 若 $A^2 + \alpha^2 I$ 非奇异, 则总可以选择初始向量 B 使得方程组有一组具有给定形式的解.

8. (a) 求下述非齐次方程组的一组特解.

$$\begin{cases} y_1' = y_1 + 3y_2 + 4\sin 2t \\ y_2' = y_1 - y_2. \end{cases}$$

(b) 求上述方程组的一组解, 使得 $y_1(0) = y_2(0) = 1$.

在习题 9 和习题 10 中, 分别求非齐次方程组 $Y'(t) = AY(t) + Q(t)$ 在给定初始条件下的解.

9. $A = \begin{bmatrix} 4 & 1 \\ -2 & 1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} 0 \\ -2e^t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

10. $A = \begin{bmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \quad Q(t) = \begin{bmatrix} t^2 \\ 2t \\ t \end{bmatrix}, \quad Y(0) = \begin{bmatrix} 6 \\ -2 \\ 1 \end{bmatrix}.$

9.17 一般线性微分方程组 $Y'(t) = P(t)Y(t) + Q(t)$

定理 9.13 给出了线性方程组

$$Y'(t) = AY(t) + Q(t), \quad Y(a) = B,$$

解的显式表示, 其中 A 是 $n \times n$ 常数矩阵而 $Q(t)$ 和 $Y(t)$ 是 $n \times m$ 矩阵函数. 现在我们来研究更一般的情形

$$Y'(t) = P(t)Y(t) + Q(t), \quad Y(a) = B, \tag{9.43}$$

其中 $n \times n$ 矩阵 $P(t)$ 不限定是常数矩阵.

如果 P 和 Q 在开区间 J 上连续, 我们将在第 10 章中证明的一个一般的存在唯一性定理告诉我们, 对 J 中任意 a 和任意初始矩阵 B, 初值问题 (9.43) 都恰有唯一的一个解. 在本节中我们将利用这个结果获得一般非齐次线性方程组解的显式表示, 从而推广了定理 9.13 的结论.

在纯量情形 ($m = n = 1$) 下, 可按下述方法求解微分方程 (9.43). 引入函数 $A(x) = \int_a^x P(t)\mathrm{d}t$, 再用 $e^{-A(t)}$ 乘公式 (9.43) 两边从而将微分方程化为如下形式:

$$e^{-A(t)}\{Y'(t) - P(t)Y(t)\} = e^{-A(t)}Q(t) \tag{9.44}$$

公式 (9.44) 的左边是函数乘积 $e^{-A(t)}Y(t)$ 的导数. 因此我们可对公式 (9.44) 的两边从 a 到 x 积分, 其中 a 和 x 为 J 中的点, 此时我们得到

$$e^{-A(x)}Y(x) - e^{-A(a)}Y(a)\} = \int_a^x e^{-A(t)}Q(t)\mathrm{d}t.$$

用 $e^{A(x)}$ 乘上式两边, 我们可得解的显式表示

$$\boldsymbol{Y}(x)=\mathrm{e}^{\boldsymbol{A}(x)}\mathrm{e}^{-\boldsymbol{A}(a)}\boldsymbol{Y}(a)+\mathrm{e}^{\boldsymbol{A}(x)}\int_a^x \mathrm{e}^{-\boldsymbol{A}(t)}\boldsymbol{Q}(t)\mathrm{d}t. \tag{9.45}$$

上述论证中唯一无法直接应用到矩阵函数的一点是 (9.44) 的左边是函数积 $\mathrm{e}^{-\boldsymbol{A}(t)}\boldsymbol{Y}(t)$ 的导数这一结论. 在这里我们用到了 $\mathrm{e}^{-\boldsymbol{A}(t)}$ 的导数等于 $-\boldsymbol{P}(t)\mathrm{e}^{-\boldsymbol{A}(t)}$ 这一事实. 在纯量情形下, 由指数函数的微分性质

$$\text{若 } \boldsymbol{E}(t)=\mathrm{e}^{\boldsymbol{A}(t)} \quad \text{则} \quad \boldsymbol{E}'(t)=\boldsymbol{A}'(t)\mathrm{e}^{\boldsymbol{A}(t)}.$$

易得该结论. 不幸的是当 $\boldsymbol{A}$ 是矩阵函数时这个微分公式并不总是成立. 例如, 该公式对 2×2 矩阵函数 $\boldsymbol{A}(t)=\begin{bmatrix} 1 & t \\ 0 & 0 \end{bmatrix}$ 不成立 (见 9.11 节习题 10). 因此我们必须对公式 (9.45) 在纯量情形下的证明加以修改, 才能将公式 (9.45) 推广到矩阵的情形.

假定我们用待定 $n\times n$ 矩阵 $\boldsymbol{F}(t)$ 乘公式 (9.43) 两边, 那么我们得到下述关系式:

$$\boldsymbol{F}(t)\boldsymbol{Y}'(t)=\boldsymbol{F}(t)\boldsymbol{P}(t)\boldsymbol{Y}(t)+\boldsymbol{F}(t)\boldsymbol{Q}(t).$$

接下来我们用 $\boldsymbol{F}'(t)\boldsymbol{Y}(t)$ 加到上式两边, 从而使等式左边变为函数积 $\boldsymbol{F}(t)\boldsymbol{Y}(t)$ 的导数, 这样我们得到

$$\{\boldsymbol{F}(t)\boldsymbol{Y}(t)\}'=\{\boldsymbol{F}'(t)+\boldsymbol{F}(t)\boldsymbol{P}(t)\}\boldsymbol{Y}(t)+\boldsymbol{F}(t)\boldsymbol{Q}(t).$$

如果我们能选取矩阵 $\boldsymbol{F}(t)$ 使得上式右边中的和式 $\{\boldsymbol{F}'(t)+\boldsymbol{F}(t)\boldsymbol{P}(t)\}$ 为零矩阵, 那么上式可简化为

$$\{\boldsymbol{F}(t)\boldsymbol{Y}(t)\}'=\boldsymbol{F}(t)\boldsymbol{Q}(t).$$

对上式两边从 a 到 x 积分可得

$$\boldsymbol{F}(x)\boldsymbol{Y}(x)-\boldsymbol{F}(a)\boldsymbol{Y}(a)=\int_a^x \boldsymbol{F}(t)\boldsymbol{Q}(t)\mathrm{d}t.$$

进一步, 如果矩阵 $\boldsymbol{F}(x)$ 非奇异, 那么我们得到显式表示

$$\boldsymbol{Y}(x)=\boldsymbol{F}(x)^{-1}\boldsymbol{F}(a)\boldsymbol{Y}(a)+\boldsymbol{F}(x)^{-1}\int_a^x \boldsymbol{F}(t)\boldsymbol{Q}(t)\mathrm{d}t. \tag{9.46}$$

这是纯量公式 (9.45) 的一个推广. 若我们能找到满足矩阵微分方程

$$\boldsymbol{F}'(t)=-\boldsymbol{F}(t)\boldsymbol{P}(t)$$

的一个 $n\times n$ 非奇异矩阵函数 $\boldsymbol{F}(t)$, 那么上述推导过程成立.

注意当 $\boldsymbol{Q}(t)=O$ 时, 上面的微分方程与原来的微分方程 (9.43) 类似, 它们的区别在于这里未知函数 $\boldsymbol{F}(t)$ 是一个方阵而不是一个 $n\times m$ 矩阵, 以及我们用 $-\boldsymbol{P}(t)$ 右乘未知函数 $\boldsymbol{F}(t)$ 而不是用 $\boldsymbol{P}(t)$ 左乘它.

接下来我们将证明上述关于 $\boldsymbol{F}$ 的微分方程总有一个非奇异解, 该证明需要用到下述齐次线性方程组的解的存在性定理:

定理 9.14　*设 $\boldsymbol{A}(t)$ 是在开区间 J 上连续的 $n\times n$ 矩阵函数, 如果 $a\in J$ 且 $\boldsymbol{B}$ 是 $n\times m$ 常数矩阵, 那么齐次线性方程组*

$$\boldsymbol{Y}'(t)=\boldsymbol{A}(t)\boldsymbol{Y}(t),\quad \boldsymbol{Y}(a)=\boldsymbol{B}$$

在 J 上有一个 $n\times m$ 矩阵解 $\boldsymbol{Y}$. □

定理 9.14 的证明将在第 10 章中给出 (定理 10.3). 借助这个结果, 我们可以证明下述定理.

定理 9.15 给定在开区间 J 上连续的 $n\times n$ 矩阵函数 $\boldsymbol{P}$ 和 J 中任意点 a, 则存在一个在 J 上满足矩阵微分方程

$$F'(x)=-F(x)P(x) \tag{9.47}$$

的 $n\times n$ 矩阵函数 $\boldsymbol{F}$ 且它的初始值 $\boldsymbol{F}(a)=\boldsymbol{I}$. 更进一步, 对 J 中任意 x, $\boldsymbol{F}(x)$ 都是非奇异的.

证明 令 $Y_k(x)$ 为微分方程

$$\boldsymbol{Y}_k'(x)=-\boldsymbol{P}(x)^{\mathrm{T}}\boldsymbol{Y}_k(x)$$

在 J 上的一个 $n\times 1$ 列矩阵解且它的初始向量为 $\boldsymbol{Y}_k(a)=\boldsymbol{I}_k$, 其中 $\boldsymbol{I}_k$ 为 $n\times n$ 单位矩阵 $\boldsymbol{I}$ 的第 k 列, $\boldsymbol{P}(x)^{\mathrm{T}}$ 为 $\boldsymbol{P}(x)$ 的转置. 再令 $\boldsymbol{G}(x)$ 为第 k 列等于 $\boldsymbol{Y}_k(x)$ 的 $n\times n$ 矩阵, 则 $\boldsymbol{G}$ 在 J 上满足矩阵微分方程

$$\boldsymbol{G}'(x)=-\boldsymbol{P}(x)^{\mathrm{T}}\boldsymbol{G}(x) \tag{9.48}$$

且它的初始值 $\boldsymbol{G}(a)=\boldsymbol{I}$. 再对等式 (9.48) 两边的矩阵取转置, 由于两个矩阵之积的转置等于这两个矩阵的转置的逆序积, 所以

$$\{\boldsymbol{G}'(x)\}^{\mathrm{T}}=-\boldsymbol{G}(x)^{\mathrm{T}}\boldsymbol{P}(x).$$

又因为 $\boldsymbol{G}$ 的导数 $\boldsymbol{G}'$ 的转置等于 $\boldsymbol{G}$ 的转置 $\boldsymbol{G}^{\mathrm{T}}$ 的导数, 所以矩阵 $\boldsymbol{F}(x)=\boldsymbol{G}(x)^{\mathrm{T}}$ 在 J 上满足微分方程 (9.47) 且它的初始值 $\boldsymbol{F}(a)=\boldsymbol{I}$.

接下来我们通过求出 $\boldsymbol{F}(x)$ 的逆来证明 $\boldsymbol{F}(x)$ 非奇异. 令 $\boldsymbol{H}$ 为一个 $n\times n$ 矩阵函数, 它的第 k 列为微分方程

$$\boldsymbol{H}'(x)=\boldsymbol{P}(x)\boldsymbol{H}(x),\quad \boldsymbol{H}(a)=\boldsymbol{I},$$

在 J 上的解. 那么对于 J 中任意 x, 乘积函数 $\boldsymbol{F}(x)\boldsymbol{H}(x)$ 的导数为

$$\boldsymbol{F}(x)\boldsymbol{H}'(x)+\boldsymbol{F}'(x)\boldsymbol{H}(x)=\boldsymbol{F}(x)\boldsymbol{P}(x)\boldsymbol{H}(x)-\boldsymbol{F}(x)\boldsymbol{P}(x)\boldsymbol{H}(x)=\boldsymbol{O}.$$

因此函数 $\boldsymbol{F}(x)\boldsymbol{H}(x)$ 为常数, 所以 $\boldsymbol{F}(x)\boldsymbol{H}(x)=\boldsymbol{F}(a)\boldsymbol{H}(a)=\boldsymbol{I}$, 由此可得 $\boldsymbol{H}(x)$ 是 $\boldsymbol{F}(x)$ 的逆. 定理得证. □

我们可将本节的结果总结为如下定理.

定理 9.16 给定 $n\times n$ 矩阵函数 $\boldsymbol{P}$ 和 $n\times m$ 矩阵函数 $\boldsymbol{Q}$, 且 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 都在开区间 J 上连续, 则初值问题

$$\boldsymbol{Y}'(x)=\boldsymbol{P}(x)\boldsymbol{Y}(x)+\boldsymbol{Q}(x),\quad \boldsymbol{Y}(a)=\boldsymbol{B}, \tag{9.49}$$

在 J 上的解由公式

$$\boldsymbol{Y}(x)=\boldsymbol{F}(x)^{-1}(a)+\boldsymbol{F}(x)^{-1}\int_a^x \boldsymbol{F}(t)\boldsymbol{Q}(t)\mathrm{d}t. \tag{9.50}$$

给出. $n\times n$ 矩阵 $\boldsymbol{F}(x)$ 的转置的第 k 列为初值问题

$$\boldsymbol{Y}'(x) = -\boldsymbol{P}(x)^{\mathrm{T}}\boldsymbol{Y}(x), \quad \boldsymbol{Y}(a) = \boldsymbol{I}_k \tag{9.51}$$

的解, 其中 $\boldsymbol{I}_k$ 为单位矩阵 $\boldsymbol{I}$ 的第 k 列. □

我们再一次提醒读者, 定理 9.16 的证明是以定理 9.14 为基础的, 但是我们尚未证明定理 9.14, 即尚未证明齐次线性微分方程组的解的存在性定理.

尽管定理 9.16 为一般线性微分方程组 (9.49) 的解提供了一个显式表示, 然而由于求矩阵函数 $\boldsymbol{F}$ 可能比较复杂, 所以在求解线性微分方程组时定理 9.16 给出的公式并非总是有用的.

9.18　求解齐次线性方程组的幂级数方法

考虑齐次线性方程组

$$Y'(x) = \boldsymbol{A}(x)\boldsymbol{Y}(x), \quad \boldsymbol{Y}(0) = \boldsymbol{B}, \tag{9.52}$$

其中给定的 $n \times n$ 矩阵 $\boldsymbol{A}(x)$ 有在某个包含原点的开区间上收敛的关于 x 的幂级数展开式, 设该幂级数展开式为

$$\boldsymbol{A}(x) = \boldsymbol{A}_0 + x\boldsymbol{A}_1 + x^2\boldsymbol{A}_2 + \cdots + x^k\boldsymbol{A}_k + \cdots, \quad |x| < r_1,$$

其中系数 $\boldsymbol{A}_0, \boldsymbol{A}_1, \boldsymbol{A}_2, \cdots$ 都是给定的 $n \times n$ 矩阵. 现在我们试图寻找方程组 (9.52) 的一组形如

$$\boldsymbol{Y}(x) = \boldsymbol{B}_0 + x\boldsymbol{B}_1 + x^2\boldsymbol{B}_2 + \cdots + x^k\boldsymbol{B}_k + \cdots,$$

的幂级数解, 其中 $\boldsymbol{B}_0, \boldsymbol{B}_1, \boldsymbol{B}_2, \cdots$, 为 $n \times m$ 常数矩阵系数. 由于 $\boldsymbol{Y}(0) = \boldsymbol{B}_0$, 所以当我们取 $\boldsymbol{B}_0 = \boldsymbol{B}$ 为预先给定的 $n \times m$ 初始矩阵时, 上式满足初始条件. 为了求其他系数, 我们将 $\boldsymbol{Y}(x)$ 和 $\boldsymbol{A}(x)$ 的幂级数展开式代入微分方程 (9.52) 并比较等号两边 x 的同次幂的系数可得如下方程组:

$$\boldsymbol{B}_1 = \boldsymbol{A}_0\boldsymbol{B}, \quad (k+1)\boldsymbol{B}_{k+1} = \sum_{r=0}^{k} \boldsymbol{A}_r\boldsymbol{B}_{k-r}, \quad k = 1, 2, \cdots \tag{9.53}$$

由上述方程组可依次求得系数矩阵 $\boldsymbol{B}_1, \boldsymbol{B}_2, \cdots$. 如果解得的 $\boldsymbol{Y}(x)$ 的幂级数在某个区间 $|x| < r_2$ 上收敛, 那么 $\boldsymbol{Y}(x)$ 为初值问题 (9.52) 在区间 $|x| < r$ 上的一组解, 其中 $r = \min\{r_1, r_2\}$.

例如, 如果 $\boldsymbol{A}(x)$ 是一个常数矩阵 $\boldsymbol{A}$, 那么 $\boldsymbol{A}_0 = \boldsymbol{A}$ 且对任意 $k \geqslant 1$ 都有 $\boldsymbol{A}_k = \boldsymbol{O}$, 所以方程组 (9.53) 化为

$$\boldsymbol{B}_1 = \boldsymbol{A}\boldsymbol{B}, \quad (k+1)\boldsymbol{B}_{k+1} = \boldsymbol{A}\boldsymbol{B}_k, \quad k \geqslant 1.$$

逐次求解该方程组可得

$$\boldsymbol{B}_k = \frac{1}{k!}\boldsymbol{A}^k\boldsymbol{B}, \quad k \geqslant 1.$$

因此在这种情形下的幂级数解为

$$Y(x) = B + \sum_{k=1}^{\infty} \frac{x^k}{k!} A^k B = \mathrm{e}^{xA} B.$$

这与在定理 9.5 中所得的常系数齐次线性微分方程组的解一致.

9.19 习　题

1. 令 p 为实值函数, Q 为 $n \times 1$ 矩阵函数, 且 p 和 Q 都在开区间 J 上连续, 再令 A 为 $n \times n$ 常数矩阵, 证明初值问题

$$Y'(x) = p(x)AY(x) + Q(x), \quad Y(a) = B,$$

在 J 上有解

$$Y(x) = \mathrm{e}^{q(x)A} B + \mathrm{e}^{q(x)A} \int_a^x \mathrm{e}^{-q(t)A} Q(t) \mathrm{d}t$$

其中 $q(x) = \int_a^x p(t)\mathrm{d}t$.

2. 考虑习题 1 的一种特殊情形: 设 A 为非奇异矩阵, $a = 0$, $p(x) = 2x$ 以及 $Q(x) = xC$, 其中 C 为常数向量. 证明此时初值问题的解变为

$$Y(x) = \mathrm{e}^{x^2 A} \left(B + \frac{1}{2} A^{-1} C \right) - \frac{1}{2} A^{-1} C.$$

3. 令 $A(t)$ 为 $n \times n$ 矩阵函数并令 $E(t) = \mathrm{e}^{A(t)}$, 再令 $Q(t)$, $Y(t)$ 和 B 为 $n \times 1$ 列矩阵. 假定在开区间 J 上 $E'(t) = A'(t)E(t)$. 如果 $a \in J$ 且 A' 和 Q 在 J 上连续, 证明初值问题 $Y'(t) = A'(t)Y(t) + Q(t)$, $Y(a) = B$ 在 J 上有解

$$Y(x) = \mathrm{e}^{A(x)} \mathrm{e}^{-A(a)} B + \mathrm{e}^{A(x)} \int_a^x \mathrm{e}^{-A(t)} Q(t) \mathrm{d}t.$$

4. 令 $E(t) = \mathrm{e}^{A(t)}$, 本题给出使得 $E'(t) = A'(t)E(t)$ 的矩阵函数 $A(t)$ 的例子.
   (a) 令 $A(t) = t^r A$, 其中 A 为 $n \times n$ 常数矩阵, r 为正整数. 证明对任意实数 t 都有 $E'(t) = A'(t)E(t)$.
   (b) 令 $A(t)$ 为 t 的多项式且该多项式的系数矩阵可交换, 即设

$$A(t) = \sum_{r=0}^{m} t^r A_r,$$

   其中对任意 r 和 s 都有 $A_r A_s = A_s A_r$. 证明对任意实数 t 都有 $E'(t) = A'(t)E(t)$.
   (c) 在 $(-\infty, +\infty)$ 上解齐次线性微分方程组 $Y'(t) = (I + tA)Y(t)$, $Y(0) = B$, 其中 A 为 $n \times n$ 常数矩阵.

5. 假定 $n \times n$ 矩阵函数 $A(x)$ 有在 $|x| < r$ 上收敛的幂级数展开式, 为二阶齐次线性方程组

$$Y(x) = A(x)Y(x), \quad Y(0) = B, \quad Y'(0) = C.$$

构造一个幂级数求解程序.

6. 考虑二阶微分方程组 $Y''(x) + AY(x) = O$, 且 $Y(0) = B$, $Y'(0) = C$, 其中 A 为 $n \times n$ 常数矩阵, 证明该方程组有对于任意实数 x 都收敛的幂级数解

$$Y(x) = \left(I + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k} A^k}{(2k)!} \right) B + \left(xI + \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k+1} A^k}{(2k+1)!} \right) C,$$

第10章　逐次逼近法

10.1　引　　言

在第 8 章和第 9 章中, 我们用到过微分方程解的存在唯一性定理, 在本章中, 我们要用**逐次逼近法**给出存在唯一性定理的证明. 逐次逼近法是一种有力的**迭代方法**(iterative procedure), 它在许多其他问题中也很有用. 逐次逼近法最初是由 Joseph Liouville 于 1838 年在关于二阶线性微分方程的研究中发表的. 随后, J. Caqué 在 1864 年, L. Fuchs 在 1870 年与 G. Peano 在 1888 年将其推广到用于 n 阶线性微分方程的研究. 1890 年, Émile Picard(1856—1941) 进一步将其推广到非线性微分方程的情形. 为纪念 Picard 的基础性贡献, 有的作者将这个方法称作 Picard 方法. 逐次逼近法不但有理论上的意义, 而且在某些情形下可用来得到解的数值逼近.

10.2　在齐次线性方程组 $\boldsymbol{Y}'(t)=\boldsymbol{A}(t)\boldsymbol{Y}(t)$ 中的应用

为介绍逐次逼近法, 我们先用它来证明下述微分方程初值问题解的存在性与唯一性:

$$\boldsymbol{Y}'(t)=\boldsymbol{A}(t)\boldsymbol{Y}(t),\quad \boldsymbol{Y}(a)=\boldsymbol{B}. \tag{10.1}$$

此处 $\boldsymbol{A}(t)$ 是一个已知的 $n\times n$ 矩阵函数, $\boldsymbol{A}(t)$ 在开区间 J 上连续, a 是 J 中任意一点. 未知函数 $\boldsymbol{Y}$ 是一个 $n\times m$ 矩阵函数, $\boldsymbol{B}$ 是一个任意给定的 $n\times m$ 常数矩阵. 在很多应用中, 我们对 $m=1$ 的情形感兴趣, 这时 $\boldsymbol{B}$ 与 $\boldsymbol{Y}$ 都是 $n\times 1$ 列矩阵, 不过这里所用的方法对 $\boldsymbol{B}$ 与 $\boldsymbol{Y}$ 为 $n\times m$ 矩阵的一般情形同样有效.

逐次逼近法的基本思想非常简单. 开始时, 我们任取一个 $\boldsymbol{Y}_0(t)$ 作为近似解, 并用它代替微分方程 (10.1) 右边的 $\boldsymbol{Y}(t)$, 由此得到一个新的微分方程

$$\boldsymbol{Y}'(t)=\boldsymbol{A}(t)\boldsymbol{Y}_0(t), \tag{10.2}$$

方程 (10.2) 的右边不再包含未知函数. 通常取初始向量 $\boldsymbol{B}$ 作为 $\boldsymbol{Y}_0(t)$, 不过这不是本质的. 方程 (10.2) 在 J 上有唯一的一个满足初始条件 $\boldsymbol{Y}_1(a)=\boldsymbol{B}$ 的解 $\boldsymbol{Y}_1$, 即

$$\boldsymbol{Y}_1(x)=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_0(t)\mathrm{d}t.$$

然后再将方程 (10.1) 右边的 $\boldsymbol{Y}(t)$ 用 $\boldsymbol{Y}_1(t)$ 来代替并再次从 a 到 x 求积, 所得的新解记作 $\boldsymbol{Y}_2$, 则 $\boldsymbol{Y}_2(a)=\boldsymbol{B}$. 换言之, 我们取

$$\boldsymbol{Y}_2(x) = \boldsymbol{B} + \int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_1(t)\mathrm{d}t.$$

重复上述步骤以至无穷, 则得到一个函数序列 $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots$, 其中 $\boldsymbol{Y}_{k+1}$ 通过递推公式

$$\boldsymbol{Y}_{k+1}(x) = \boldsymbol{B} + \int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_k(t)\mathrm{d}t, \quad k = 0, 1, 2, \cdots \tag{10.3}$$

从 $\boldsymbol{Y}_k$ 求出.

一个值得指出的事实是, 这样得到的函数序列一定收敛于一个极限函数 $\boldsymbol{Y}$, 而 $\boldsymbol{Y}$ 是初值问题 (10.1) 的解且是它的唯一解. 在证明上述程序的收敛性之前, 我们先给出一个例子来说明我们所用的方法, 这个例子的解我们已用别的方法求得.

例 1　考虑初值问题 $\boldsymbol{Y}'(t) = \boldsymbol{A}\boldsymbol{Y}(t), \boldsymbol{Y}(0) = \boldsymbol{B}$, 此处 $\boldsymbol{A}$ 是一个 $n \times n$ 常数矩阵. 我们由定理 9.7 知道这个问题的解由公式 $Y(x) = \mathrm{e}^{x\boldsymbol{A}}\boldsymbol{B}$ 给出, 此处 x 为任意实数. 为了用逐次逼近法得到这个解, 我们取 $\boldsymbol{Y}_0(x) = \boldsymbol{B}$. 重复应用递推公式 (10.3) 得

$$\begin{aligned}
\boldsymbol{Y}_1(x) &= \boldsymbol{B} + \int_0^x \boldsymbol{A}\boldsymbol{B}\mathrm{d}t = \boldsymbol{B} + x\boldsymbol{A}\boldsymbol{B}, \\
\boldsymbol{Y}_2(x) &= \boldsymbol{B} + \int_0^x \boldsymbol{A}\boldsymbol{Y}_1(t)\mathrm{d}t = \boldsymbol{B} + \int_0^x (\boldsymbol{A}\boldsymbol{B} + t\boldsymbol{A}^2\boldsymbol{B})\mathrm{d}t = \boldsymbol{B} + x\boldsymbol{A}\boldsymbol{B} + \frac{1}{2}x^2\boldsymbol{A}^2\boldsymbol{B}, \\
&\ \ \vdots \\
\boldsymbol{Y}_k(x) &= \boldsymbol{B} + x\boldsymbol{A}\boldsymbol{B} + \frac{1}{2}x^2\boldsymbol{A}^2\boldsymbol{B} + \cdots + \frac{1}{k!}x^k\boldsymbol{A}^k\boldsymbol{B} = \left(\sum_{r=0}^{k} \frac{(x\boldsymbol{A})^r}{r!}\right)\boldsymbol{B}.
\end{aligned}$$

最后一式右边的和是级数 $\mathrm{e}^{x\boldsymbol{A}}$ 的一个部分和. 因此当 $k \to \infty$ 时我们有

$$\lim_{k\to\infty} \boldsymbol{Y}_k(x) = \mathrm{e}^{x\boldsymbol{A}}\boldsymbol{B}.$$

□

10.3　逐次逼近序列的收敛性

定理 10.1 (齐次线性微分方程组解的存在性定理)　*给定一个在开区间 J 上连续的 $n \times n$ 矩阵函数 $\boldsymbol{A}(t)$, 给定任意 $n \times m$ 常数矩阵 $\boldsymbol{B}$. 若 $a \in J$, 则由递推公式*

$$\boldsymbol{Y}_0(x) = \boldsymbol{B}, \quad \boldsymbol{Y}_{k+1}(x) = \boldsymbol{B} + \int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_k(t)\mathrm{d}t, \quad k = 0, 1, 2, \cdots \tag{10.4}$$

定义 $n \times m$ 矩阵函数序列 $\boldsymbol{Y}_0(x), \boldsymbol{Y}_1(x), \boldsymbol{Y}_2(x), \cdots,$. 于是对 J 中的每个 x, 序列 $\{\boldsymbol{Y}_k(x)\}$ 收敛到一个满足初值问题

$$\boldsymbol{Y}'(x) = \boldsymbol{A}(x)\boldsymbol{Y}(x), \quad \boldsymbol{Y}(a) = \boldsymbol{B} \tag{10.5}$$

的 $n \times m$ 矩阵函数 $\boldsymbol{Y}(x)$.

证明　首先, 我们将各项 $\boldsymbol{Y}_k(x)$ 表示成如下形式:

$$\boldsymbol{Y}_k(x)=\boldsymbol{Y}_0(x)+\sum_{r=0}^{k-1}\{\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\}. \tag{10.6}$$

为证明当 $k\to\infty$ 时, $\boldsymbol{Y}_k(x)$ 趋于一个极限, 我们要证明对 J 中的每个 x, 无穷级数

$$\sum_{r=0}^{\infty}\{\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\} \tag{10.7}$$

都收敛. 为此只需证明矩阵范数的级数

$$\sum_{r=0}^{\infty}\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\| \tag{10.8}$$

收敛即可. 在这个级数中我们用的是 9.3 节中介绍的矩阵范数, 即一个矩阵的范数是它的所有元素的绝对值之和.

考虑 J 的一个包含 a 的有界闭区间 J_1, 我们要证明对 J_1 中的每一个 x, 级数 (10.8) 都被一个与 x 无关的收敛数项级数所控制. 这意味着级数 (10.7) 不但收敛, 而且在 J_1 上一致收敛.

为了估计级数 (10.8) 中各项的大小, 我们反复使用递推公式. 作为开始, 我们有

$$\boldsymbol{Y}_1(x)-\boldsymbol{Y}_0(x)=\int_a^x\boldsymbol{A}(t)\boldsymbol{B}\mathrm{d}t.$$

为简便计, 我们假定 $a<x$, 于是得

$$\|\boldsymbol{Y}_1(x)-\boldsymbol{Y}_0(x)\|=\left\|\int_a^x\boldsymbol{A}(t)\boldsymbol{B}\mathrm{d}t\right\|\leqslant\int_a^x\|\boldsymbol{A}(t)\|\|\boldsymbol{B}\|\mathrm{d}t. \tag{10.9}$$

由于 $\boldsymbol{A}(t)$ 的各元素在 J 上连续, 因此各元素在有界闭子区间 J_1 上有界. 从而 $\|\boldsymbol{A}(t)\|\leqslant M$, 此处 M 是 $\boldsymbol{A}(t)$ 的所有元素在区间 J_1 上的界的和 (数 M 与 J_1 有关). 从而不等式 (10.9) 中的被积函数的绝对值不超过 $\|\boldsymbol{B}\|M$, 因此对 J_1 中所有 $x>a$, 都有

$$\|\boldsymbol{Y}_1(x)-\boldsymbol{Y}_0(x)\|\leqslant\int_a^x\|\boldsymbol{B}\|M\mathrm{d}t=\|\boldsymbol{B}\|M(x-a).$$

再次利用递推公式可将差 $\boldsymbol{Y}_2-\boldsymbol{Y}_1$ 用 $\boldsymbol{Y}_1-\boldsymbol{Y}_0$ 表示, 于是由刚才得到的关于 $\boldsymbol{Y}_1-\boldsymbol{Y}_0$ 的估计, 对 J_1 中所有的 $x>a$, 都有

$$\begin{aligned}\|\boldsymbol{Y}_2(x)-\boldsymbol{Y}_1(x)\|&=\left\|\int_a^x\boldsymbol{A}(t)\{\boldsymbol{Y}_1(t)-\boldsymbol{Y}_0(t)\}\mathrm{d}t\right\|\leqslant\int_a^x\|\boldsymbol{A}(t)\|\|\boldsymbol{B}\|M(t-a)\mathrm{d}t\\&\leqslant\|\boldsymbol{B}\|M^2\int_a^x(t-a)\mathrm{d}t=\|\boldsymbol{B}\|\frac{M^2(x-a)^2}{2!}.\end{aligned}$$

再由归纳法即得, 对 J_1 中所有的 $x>a$, 都有

$$\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\|\leqslant\|\boldsymbol{B}\|\frac{M^{r+1}(x-a)^{r+1}}{(r+1)!},\quad r=0,1,2,\cdots.$$

若 $x < a$, 由类似的讨论可得同样的不等式, 只不过需要将 $(x-a)$ 用 $|x-a|$ 代替而已. 如果我们用 L 表示区间 J_1 的长, 则对 J_1 中所有 x, 都有 $|x-a| \leqslant L$. 因此对 J_1 中所有的 x, 都有

$$\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\| \leqslant \|\boldsymbol{B}\|\frac{M^{r+1}L^{r+1}}{(r+1)!},\quad r=0,1,2,\cdots.$$

因此级数 (10.8) 被收敛级数

$$\|\boldsymbol{B}\|\sum_{r=0}^{\infty}\frac{(ML)^{r+1}}{(r+1)!}=\|\boldsymbol{B}\|(\mathrm{e}^{ML}-1)$$

所控制. 由此证明了级数 (10.8) 在 J_1 上一致收敛. □

上面的讨论表明, 逐次逼近序列一定收敛而且在 J_1 上一致收敛. 令 $\boldsymbol{Y}$ 表示这个收敛序列的极限函数, 即对 J_1 中的每一个 x, 由

$$\boldsymbol{Y}(x)=\lim_{k\to\infty}\boldsymbol{Y}_k(x)$$

定义函数 $\boldsymbol{Y}(x)$, 我们将证明 $\boldsymbol{Y}$ 具有下列性质.

(a) $\boldsymbol{Y}$ 在 J_1 上连续.

(b) $\boldsymbol{Y}(x)=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}(t)\mathrm{d}t$, 对所有 $x\in J_1$.

(c) $\boldsymbol{Y}(a)=\boldsymbol{B}$ 且 $\boldsymbol{Y}'(x)=\boldsymbol{A}(x)\boldsymbol{Y}(x)$, 对所有 $x\in J_1$.

性质 (c) 指出, $\boldsymbol{Y}$ 是 J_1 上初值问题的一个解.

(a) 的证明 每个函数 $\boldsymbol{Y}_k$ 都是一个 $n\times m$ 矩阵, 它的元素都是在 J_1 上连续的纯量函数. 极限函数 $\boldsymbol{Y}$ 的元素都是 J_1 上连续函数的一致收敛序列的极限, 因此 $\boldsymbol{Y}$ 的各个元素也都是 J_1 上的连续函数, 从而 $\boldsymbol{Y}$ 本身也在 J_1 上连续.

(b) 的证明 由递推公式 (10.4), 我们有

$$\boldsymbol{Y}_{k+1}(x)=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_k(t)\mathrm{d}t.$$

由于序列 $\{Y_k\}$ 在 J_1 上一致收敛, 因此极限符号与积分号可以互换, 即

$$\begin{aligned}\boldsymbol{Y}(x)&=\lim_{k\to\infty}\boldsymbol{Y}_{k+1}(x)=\boldsymbol{B}+\lim_{k\to\infty}\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_k(t)\mathrm{d}t\\&=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\lim_{k\to\infty}\boldsymbol{Y}_k(t)\mathrm{d}t=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}(t)\mathrm{d}t.\end{aligned}$$

(c) 的证明 由 (b) 立得 $\boldsymbol{Y}(a)=\boldsymbol{B}$. 由性质 (a) 可知, (b) 中的被积函数在 J_1 上连续, 因此由微积分第一基本定理得 $\boldsymbol{Y}'(x)$ 存在且在 J_1 上等于 $\boldsymbol{A}(x)\boldsymbol{Y}(x)$.

J_1 可以是 J 中包含 a 的任意一个有界闭子区间. 若将 J_1 放大, 求 $\boldsymbol{Y}(x)$ 的过程不变, 因为它只牵涉到从 a 到 x 的积分. 对 J 中的每一个 x 都存在 J 的一个同时包含 a 与 x 的有界闭子区间. 从而初值问题在全区间 J 上存在一个解. □

定理 10.2 (齐次线性微分方程组解的唯一性定理) 设 $\boldsymbol{A}(t)$ 在开区间 J 上连续, 则微分方程组

$$\boldsymbol{Y}'(t) = \boldsymbol{A}(t)\boldsymbol{Y}(t)$$

在 J 上至多存在一个满足给定初始条件 $\boldsymbol{Y}(a) = \boldsymbol{B}$ 的解.

证明　设 $\boldsymbol{Y}$ 与 $\boldsymbol{Z}$ 为在 J 上的两个解. 令 J_1 为 J 的包含 a 的任一有界闭子区间. 我们要证明对 J_1 中任意一个 x 都有 $\boldsymbol{Z}(x) = \boldsymbol{Y}(x)$, 由此即得在全区间 J 上有 $\boldsymbol{Z} = \boldsymbol{Y}$.

由于 $\boldsymbol{Y}$ 与 $\boldsymbol{Z}$ 都是解, 因此

$$\boldsymbol{Z}'(t) - \boldsymbol{Y}'(t) = \boldsymbol{A}(t)\{\boldsymbol{Z}(t) - \boldsymbol{Y}(t)\}.$$

在 J_1 中选取 x 并对上式两边从 a 到 x 积分, 由 $\boldsymbol{Z}(a) = \boldsymbol{Y}(a)$ 得

$$\boldsymbol{Z}(x) - \boldsymbol{Y}(x) = \int_a^x \boldsymbol{A}(t)\{\boldsymbol{Z}(t) - \boldsymbol{Y}(t)\}\mathrm{d}t.$$

由此即得不等式

$$\|\boldsymbol{Z}(x) - \boldsymbol{Y}(x)\| \leqslant M\left|\int_a^x \|\boldsymbol{Z}(t) - \boldsymbol{Y}(t)\|\mathrm{d}t\right|, \tag{10.10}$$

其中 M 是 $\|\boldsymbol{A}(t)\|$ 在 J_1 上的一个上界. 令 M_1 为 J_1 上连续函数 $\|\boldsymbol{Z}(t) - \boldsymbol{Y}(t)\|$ 的一个上界. 由不等式 (10.10) 得

$$\|\boldsymbol{Z}(x) - \boldsymbol{Y}(x)\| \leqslant MM_1|x - a|. \tag{10.11}$$

将不等式 (10.11) 用于不等式 (10.10) 的右边得

$$\|\boldsymbol{Z}(x) - \boldsymbol{Y}(x)\| \leqslant M^2M_1\left|\int_a^x |t - a|\mathrm{d}t\right| = M^2M_1\frac{|x - a|^2}{2}.$$

由归纳法得, 对每一个 $r \geqslant 1$ 都有

$$\|\boldsymbol{Z}(x) - \boldsymbol{Y}(x)\| \leqslant M^rM_1\frac{|x - a|^r}{r!}. \tag{10.12}$$

当 $r \to \infty$ 时上式右边趋于 0, 因此 $\|\boldsymbol{Z}(x) - \boldsymbol{Y}(x)\| = 0$, 从而 $\boldsymbol{Z}(x) = \boldsymbol{Y}(x)$, 完成证明. □

将定理 10.1 与定理 10.2 相结合即得下述存在唯一性定理.

定理 10.3　*设 $\boldsymbol{A}(t)$ 为在开区间 J 上连续的 $n \times n$ 矩阵函数. 若 $a \in J$ 且 $\boldsymbol{B}$ 是任一 $n \times m$ 常数矩阵, 则齐次线性微分方程组*

$$\boldsymbol{Y}'(t) = \boldsymbol{A}(t)\boldsymbol{Y}(t), \quad \boldsymbol{Y}(a) = \boldsymbol{B}$$

在 J 上存在且只存在一个 $n \times m$ 矩阵解 $\boldsymbol{Y}$. □

10.4　用于一阶非线性方程组的逐次逼近法

逐次逼近法也可以应用于某些非线性微分方程组. 考虑形如

$$\boldsymbol{Y}' = \boldsymbol{F}(t, \boldsymbol{Y}) \tag{10.13}$$

的一阶微分方程组, 其中 $\boldsymbol{F}$ 是给定的 $n \times m$ 矩阵函数, $\boldsymbol{Y}$ 是待确定的 $n \times m$ 未知矩阵函数. 我们要寻找对某个区间 J 中的每一个 t 都满足方程

$$\boldsymbol{Y}'(t) = \boldsymbol{F}[t, \boldsymbol{Y}(t)]$$

并且还满足给定的初始条件 $\boldsymbol{Y}(a) = \boldsymbol{B}$ 的一个解 $\boldsymbol{Y}$, 此处 $a \in J$, $\boldsymbol{B}$ 是一个给定的 $n \times m$ 常数矩阵.

与线性的情形相似, 我们构造一个逐次逼近序列 $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots$, 其中 $\boldsymbol{Y}_0 = \boldsymbol{B}$, 对 $k = 0, 1, 2, \cdots$, $\boldsymbol{Y}_{k+1}$ 由 $\boldsymbol{Y}_k$ 通过递推公式

$$\boldsymbol{Y}_{k+1}(x) = \boldsymbol{B} + \int_a^x \boldsymbol{F}[t, \boldsymbol{Y}_k(t)]\mathrm{d}t \tag{10.14}$$

给出. 对 $\boldsymbol{F}$ 作某些限制, 这个序列将收敛于一个满足给定微分方程和给定初始条件的极限函数 $\boldsymbol{Y}$.

在研究为保证收敛性而应对 $\boldsymbol{F}$ 加上哪些限制条件之前, 我们先讨论几个具体例子以说明在实际操作时可能会遇到的某些困难.

例 1 考虑非线性初值问题 $y' = x^2 + y^2$, 当 $x = 0$ 时 $y = 0$. 我们来求逐次逼近序列的前几项. 选取 $Y_0(x) = 0$, 然后由递推公式 (10.14) 计算后面的 3 个近似解:

$$Y_1(x) = \int_0^x t^2 \mathrm{d}t = \frac{x^3}{3},$$

$$Y_2(x) = \int_0^x [t^2 + Y_1^2(t)]\mathrm{d}t = \int_0^x \left(t^2 + \frac{t^6}{9}\right)\mathrm{d}t = \frac{x^3}{3} + \frac{x^7}{63},$$

$$Y_3(x) = \int_0^x \left[t^2 + \left(\frac{t^3}{3} + \frac{t^7}{63}\right)^2\right]\mathrm{d}t = \frac{x^3}{3} + \frac{x^7}{63} + \frac{2x^{11}}{2079} + \frac{x^{15}}{59535}.$$

显然, 为计算后面的近似解, 工作量是巨大的. 例如, 后面两个近似解 Y_4 与 Y_5 将分别是 31 次与 63 次的多项式. □

下述例子指出了在计算逐次近似解时可能会遇到的更严重的困难.

例 2 考虑非线性初值问题 $y' = 2x + \mathrm{e}^y$, 此处当 $x = 0$ 时 $y = 0$. 我们取 $Y_0(x) = 0$, 由递推公式得

$$Y_1(x) = \int_0^x (2t + 1)\mathrm{d}t = x^2 + x,$$

$$Y_2(x) = \int_0^x (2t + \mathrm{e}^{t^2+t})\mathrm{d}t = x^2 + \int_0^x \mathrm{e}^{t^2+t}\mathrm{d}t.$$

由于上述最后一个积分不能用初等函数表出, 因此进一步的计算受阻. 不过, 对给定的一个 x, 我们仍然可以求出积分的近似值并从而得到 $Y_2(x)$ 的一个近似值. □

由于在上述两例中所展示的困难, 逐次逼近法对于确定解的显式表达式而言并不十分有用. 这个方法的实际价值在于用来建立存在性定理.

10.5 一阶非线性方程组解的存在唯一性定理的证明

现在我们转到关于一阶非线性方程组的一个存在唯一性定理上来. 对出现在微分方程组

$$\boldsymbol{Y}' = \boldsymbol{F}(x, \boldsymbol{Y})$$

右边的函数 F 加上适当的限制条件, 我们可以将 10.3 节中关于线性情形的证明中所用的方法加以推广.

令 J 表示我们要在它上面求解的开区间. 设 $a \in J$, 令 $\boldsymbol{B}$ 为给定的 $n \times m$ 常数矩阵. 令 S 表示由

$$S = \{(x, \boldsymbol{Y}) : |x - a| \leqslant h, \|\boldsymbol{Y} - \boldsymbol{B}\| \leqslant b\}$$

给出的序对 $(x, \boldsymbol{Y})$ 的集合, 其中 $h > 0$, $b > 0$. (若 $n = m = 1$, 这是一个以 (a, B) 为中心, 底为 $2h$, 高为 $2b$ 的矩形.) 假定 F 的定义域包含一个这样的集合 S 并且 F 在 S 上有界, 即设对 S 中所有 $(x, \boldsymbol{Y})$ 都有

$$\|\boldsymbol{F}(x, \boldsymbol{Y})\| \leqslant M, \tag{10.15}$$

其中 M 为正常数.

其次, 我们假定, 对每一个在区间 $(a - h, a + h)$ 上连续并且对所有 $x \in (a - h, a + h)$ 都有 $(x, \boldsymbol{Y}(x)) \in S$ 的函数 $\boldsymbol{Y}$, 复合矩阵函数 $\boldsymbol{G}(x) = \boldsymbol{F}[x, \boldsymbol{Y}(x)]$ 在区间 $(a - h, a + h)$ 上也连续. 这个假设保证了逐次逼近法中牵涉到的积分的存在性, 并且还保证了这样构造出的函数的连续性.

最后, 我们还假设 $\boldsymbol{F}$ 满足形如

$$\|\boldsymbol{F}(x, \boldsymbol{Y}) - \boldsymbol{F}(x, \boldsymbol{Z})\| \leqslant C\|\boldsymbol{Y} - \boldsymbol{Z}\|$$

的条件, 此处 $(x, \boldsymbol{Y}) \in S, (x, \boldsymbol{Z}) \in S$, C 是一个正常数. 这个限制条件叫作 **Lipschitz 条件**, 这是为了纪念德国数学家 Rudolph Lipschitz(1832—1903), 他在 1876 年首次引入了这个条件. Lipschitz 条件并未对函数作出过多的限制, 但能使我们把关于存在性定理的证明从线性情形推广到非线性情形.

定理 10.4 假定 $\boldsymbol{F}$ 在集合 S 上满足有界性、连续性和 Lipschitz 条件. 令 I 表示开区间 $(a - L, a + L)$, 此处 L 为 h 与 b/M 中的较小者. 则存在唯一的一个定义在 I 上的 $n \times m$ 矩阵函数 $\boldsymbol{Y}$, 使得 $\boldsymbol{Y}(a) = \boldsymbol{B}$, $(x, \boldsymbol{Y}(x)) \in S$ 并且对任意 $x \in I$ 都有

$$\boldsymbol{Y}'(x) = \boldsymbol{F}[x, \boldsymbol{Y}(x)].$$

证明 由于定理的证明与线性情形类似, 因此我们只简单叙述主要步骤. 令 $\boldsymbol{Y}_0(x) = \boldsymbol{B}$ 并由递推公式

$$\boldsymbol{Y}_{k+1}(x) = \boldsymbol{B} + \int_a^x F[t, Y_k(t)]\mathrm{d}t, \quad k = 0, 1, 2, \cdots \tag{10.16}$$

定义 I 上的 $n\times m$ 矩阵函数 $y_1, y_2, \cdots$. 为使这个递推公式有意义, 我们需要确定对每一个 $x\in I$ 都有 $(x, \boldsymbol{Y}_k(x))\in S$. 由归纳法可以很简单地证明这一点. 当 $k=0$ 时我们有 $(x, \boldsymbol{Y}_0(x))=(x, \boldsymbol{B})\in S$. 今设对某个 k 与所有 $x\in I$ 都有 $(x, \boldsymbol{Y}_k(x))\in S$. 由 (10.16) 与 (10.15) 两式得

$$\|\boldsymbol{Y}_{k+1}(x)-\boldsymbol{B}\|\leqslant\left|\int_a^x\|\boldsymbol{F}[t, \boldsymbol{Y}_k(t)]\|\mathrm{d}t\right|\leqslant M\left|\int_a^x \mathrm{d}t\right|=M|x-a|.$$

由于对 $x\in I$ 都有 $|x-a|\leqslant L$. 因此得

$$\|\boldsymbol{Y}_{k+1}(x)-\boldsymbol{B}\|\leqslant ML\leqslant b,$$

由此可知对每一个 $x\in I$ 都有 $(x, \boldsymbol{Y}_{k+1}(x))\in S$. 因此递推公式对每一个 $k\geqslant 0$ 与每一个 $x\in I$ 都有意义.

现在可以和 10.3 节完全一样来确定序列 $\{\boldsymbol{Y}_k(x)\}$ 的收敛性. 将 $\boldsymbol{Y}_k(x)$ 表成下述形式:

$$\boldsymbol{Y}_k(x)=\boldsymbol{Y}_0(x)+\sum_{r=0}^{k-1}\{\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\},$$

再通过证明无穷级数

$$\sum_{r=0}^{\infty}\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\|$$

在 I 上收敛而推出当 $k\to\infty$ 时, $\boldsymbol{Y}_k(x)$ 趋于一个极限. 利用递推公式和 Lipschitz 条件, 由归纳法可得

$$\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\|\leqslant\frac{MC^r|x-a|^{r+1}}{(r+1)!}\leqslant\frac{MC^rL^{r+1}}{(r+1)!}.$$

由此即得级数

$$\sum_{r=0}^{\infty}\|\boldsymbol{Y}_{r+1}(x)-\boldsymbol{Y}_r(x)\|$$

的收敛性. 于是, 对 $x\in I$, 可由

$$\boldsymbol{Y}(x)=\lim_{k\to\infty}\boldsymbol{Y}_k(x)$$

来定义极限函数 $\boldsymbol{Y}$. 和线性的情形完全一样, 可以验证 $\boldsymbol{Y}$ 满足积分方程

$$\boldsymbol{Y}(x)=\boldsymbol{B}+\int_a^x\boldsymbol{F}[t, \boldsymbol{Y}(t)]\mathrm{d}t.$$

由此证明了解的存在性. 再用和证明定理 10.2 相同的方法可以证明唯一性. □

10.6 习 题

1. 考虑线性初值问题 $y'+y=2\mathrm{e}^x$, 当 $x=0$ 时 $y=1$.

   (a) 求这个一阶线性微分方程的精确解 $\boldsymbol{Y}$.

(b) 应用逐次逼近法, 对 $Y_0(x)=1$, 求出 $Y_n(x)$ 的显式表达式并证明对所有实数 x 都有 $\lim_{n\to\infty} Y_n(x)=Y(x)$.

2. 将逐次逼近法用于非线性初值问题

$$y' = x + y^2, \quad 当x = 0时y = 0.$$

取 $Y_0(x) = 0$, 求 $Y_3(x)$.

3. 将逐次逼近法用于非线性初值问题

$$y' = 1 + xy^2, \quad 当x = 0时y = 0.$$

取 $Y_0(x) = 0$, 求 $Y_3(x)$.

4. 将逐次逼近法用于非线性初值问题

$$y' = x^2 + y^2, \quad 当x = 0时y = 0.$$

从一个"坏"的初始近似 $Y_0(x) = 1$ 出发求 $Y_3(x)$, 并与 10.4 节中例 1 的结果比较.

5. 设 a 与 b 为实数, 令 $M(a, b)$ 表示 a 与 b 中较大者. 考虑在区间 $[-1, 1]$ 上的下述初值问题:

$$y' = M(x, y), \quad 当x = 0时y = 1.$$

(a) 取 $Y_0(x) = 1$, 对 $n = 1, 2, 3, 4$, 用逐次逼近法求 $Y_n(x)$.

(b) 求此初值问题的解.

6. 考虑非线性初值问题:

$$y' = x^2 + y^2, \quad 当x = 0时y = 1.$$

(a) 取 $Y_0(x) = 1$, 用逐次逼近法求 $Y_2(x)$.

(b) 令 $f(x, y) = x^2 + y^2$, 考虑矩形区域 $R = [-1, 1] \times [-1, 1]$, 求最小的实数 M, 使得对所有 $(x, y) \in R$ 都有 $|f(x, y)| \leqslant M$. 找出一个区间 $I = (-c, c)$ 使得 I 上的每一个逼近函数 Y_n 的图像都落在 R 中.

(c) 假定解 $y = Y(x)$ 在原点的某个邻域内能展开成幂级数. 求这个幂级数展开式的前 6 个非零项并与 (a) 中的结果相比较.

7. 考虑非线性初值问题:

$$y' = 1 + y^2, \quad 当x = 0时y = 0.$$

(a) 取 $Y_0(x) = 0$, 用逐次逼近法求 $Y_4(x)$.

(b) 证明每一个逼近函数 Y_n 都在整个实轴上有定义.

(c) 利用定理 10.4 证明这个初值问题在任意一个形如 $(-h, h)$ 的区间内至多只能有一个解.

(d) 验证函数 $y = \tan x$ 在区间 $(-\pi/2, \pi/2)$ 中满足本题给出的微分方程. 在本题中, 各个逼近函数在整个实轴上都有定义, 但它们只在区间 $(-\pi/2, \pi/2)$ 中收敛到一个极限函数.

8. 求同时满足方程组 $y' = z, z' = x^3(y + z)$ 及初始条件 $x = 0$ 时 $y = 1$ 及 $z = 1/2$ 的两个函数 $y = Y(x)$ 与 $z = Z(x)$. 从初始近似 $Y_0(x) = 1$ 与 $Z_0(x) = 1/2$ 出发, 用逐次逼近法求出逼近函数

$$Y_3(x) = 1 + \frac{x}{2} + \frac{3x^5}{40} + \frac{x^6}{60} + \frac{x^9}{192},$$

$$Z_3(x) = \frac{1}{2} + \frac{3x^4}{8} + \frac{x^5}{10} + \frac{3x^8}{64} + \frac{7x^9}{360} + \frac{x^{12}}{256}.$$

9. 考虑方程组 $y' = 2x + z, z' = 3xy + x^2 z$, 初始条件为 $x = 0$ 时 $y = 2, z = 0$. 从初始近似 $Y_0(x) = 2, Z_0(x) = 0$ 出发, 用逐次逼近法求 $Y_3(x)$ 与 $Z_3(x)$.

10. 考虑初值问题 $y'' = x^2 y' + x^4 y$, $x = 0$ 时 $y = 5, y' = 1$. 将它转化为一个由关于两个未知函数 $y = Y(x)$ 和 $z = Z(x)$ 的两个方程组成的方程组的等价问题, 此处 $z = y'$. 然后从初始近似 $Y_0(x) = 5$ 和 $Z_0(x) = 1$ 出发用逐次逼近法求 $Y_3(x)$ 与 $Z_3(x)$.

*10.7[①] 逐次逼近与算子不动点

逐次逼近法的基本思想不但可以用来建立微分方程的存在性定理, 而且还可以用于分析其他许多重要问题. 下面几节将从一个新的视角来重述逐次逼近法, 由此可大大地扩展其应用范围.

在定理 10.1 的证明中, 我们依据递推公式

$$\boldsymbol{Y}_{k+1}(x)=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}_k(t)\mathrm{d}t$$

构造出一个函数序列 $\{\boldsymbol{Y}_k\}$. 上式的右边可以看作这样一个算子 T, 它把某些函数 $\boldsymbol{Y}$ 转化为由

$$T(\boldsymbol{Y})=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}(t)\mathrm{d}t$$

定义的新函数 $T(\boldsymbol{Y})$. 在定理 10.1 的证明中, 我们发现初值问题 $\boldsymbol{Y}'(x)=\boldsymbol{A}(x)\boldsymbol{Y}(x)$, $\boldsymbol{Y}(a)=\boldsymbol{B}$ 的解 $\boldsymbol{Y}$ 满足积分方程

$$\boldsymbol{Y}=\boldsymbol{B}+\int_a^x \boldsymbol{A}(t)\boldsymbol{Y}(t)\mathrm{d}t.$$

将上式用算子记号表示即为 $T(\boldsymbol{Y})=\boldsymbol{Y}$. 换言之, 解 $\boldsymbol{Y}$ 在算子 T 作用下不变. 这样的一个函数 $\boldsymbol{Y}$ 叫作算子 T 的一个**不动点**(fixed point).

分析中的许多问题可以用某种适当的形式提出, 使得问题的解依赖于某个算子的不动点的存在性. 现在我们就来系统地讨论这个问题.

*10.8 赋范线性空间

为了从一个一般的形式给出逐次逼近法, 在线性空间的框架下来进行研究是比较方便的. 设 S 为任意一个线性空间. 当我们论及 S 中的一个元素 x 用 S 中另一个元素 y 逼近时, 我们考虑差 $x-y$ 并将它称作逼近的**误差**(error). 为了度量误差的大小, 我们在空间中引入范数. 任一欧氏空间总有一个从内积引出的范数, 即 $\|x\|=(x,x)^{1/2}$. 不过我们也对不是从内积引出的范数感兴趣, 所以我们给出一个在一般意义下的范数的定义.

定义(范数) 设 S 为任意线性空间. 定义在 S 上的一个实值函数 N 若具有性质:

(a) 对任意 $x\in S$, 都有 $N(x)\geqslant 0$,

(b) 对任意 $x\in S$ 和任意纯量 c, 都有 $N(cx)=|c|N(x)$,

① 本章中加星号的小节是选学内容, 这些小节介绍了线性代数中的概念在分析学中的进一步应用.

(c) 对任意 $x, y \in S$, 都有 $N(x+y) \leqslant N(x) + N(y)$,

(d) 若 $N(x) = 0$, 则必定有 $x = O$,

则称 N 为一个**范数**(norm). 定义了一个范数的线性空间叫作**赋范线性空间**(normed linear space).

x 的范数 $N(x)$ 有时也记作 $\|x\|$. 若采用这个记号, 则关于范数定义的性质变为:

(a) 对所有 $x \in S$, 都有 $\|x\| \geqslant 0$;

(b) 对所有 $x \in S$ 及所有纯量 c, 都有 $\|cx\| = |c|\|x\|$;

(c) 对所有 $x, y \in S$, 都有 $\|x+y\| \leqslant \|x\| + \|y\|$;

(d) 若 $\|x\| = 0$, 则必有 $x = O$,

对 $x, y \in S$, 称 $\|x - y\|$ 为 x 到 y 的**距离**.

例(最大范数) 令 $C(J)$ 表示定义在有界闭区间 J 上的全体实值函数组成的线性空间. 对 $\varphi \in C(J)$, 定义

$$\|\varphi\| = \max_{x \in J} |\varphi(x)|,$$

上式右边表示 φ 在 J 上取得的最大绝对值. 读者可以验证它具有范数定义中的 4 条性质. 这个范数叫作**最大范数**. □

最大范数不可能从内积导出. 为证明这一点, 我们需要指出它不具有所有内积范数都具有的某些性质. 例如, 若一个范数是由内积导出的, 则对所有 $x, y \in S$, **平行四边形法则**

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

都成立 (见 3.13 节习题 8). 但对于极大范数, 平行四边形法则并不总是成立. 例如, 令 x 与 y 为区间 $[0,1]$ 上由

$$x(t) = t, \quad y(t) = 1 - t$$

给出的两个函数. 则我们有 $\|x\| = \|y\| = \|x+y\| = \|x-y\| = 1$, 因此平行四边形法则不成立.

*10.9 收缩算子

在本节中, 我们考虑由在有界闭区间 J 上的全体连续实值函数关于极大范数 $\|\varphi\|$ 所组成的赋范线性空间 $C(J)$. 考虑算子

$$T : C(J) \to C(J),$$

它的定义域是 $C(J)$, 值域是 $C(J)$ 的一个子集. 即若 φ 在 J 上连续, 则 $T(\varphi)$ 也在 J 上连续. 下列公式给出了这类算子的一些简单例子. 在各例中, φ 都是 $C(J)$ 中的

任意函数, 对每个 $x \in J$, $T(\varphi)(x)$ 则分别由下述公式给出:

$$T(\varphi)(x) = \lambda\varphi(x),\quad \lambda\text{为给定实数},$$

$$T(\varphi)(x) = \int_c^x \varphi(t)\mathrm{d}t,\quad c\text{为}J\text{中给定的一点},$$

$$T(\varphi)(x) = b + \int_c^x f[t,\varphi(t)]\mathrm{d}t,\quad \text{此处}b\text{为常数且复合函数}f[t,\varphi(t)]\text{在}J\text{上连续}.$$

我们对那些使得距离 $\|T(\varphi) - T(\psi)\|$ 小于 $\|\varphi - \psi\|$ 与某个固定的常数 α 之积的算子 T 感兴趣, 这里常数 $\alpha < 1$. 这些算子叫作**收缩算子**, 其定义如下.

定义(收缩算子) 设 $T: C(J) \to C(J)$ 为一个算子, 如果存在常数 α, 其中 $0 \leqslant \alpha < 1$, 使得对于 $C(J)$ 中的任意一对函数 φ 与 ψ, 都有

$$\|T(\varphi) - T(\psi)\| \leqslant \alpha\|\varphi - \psi\|, \tag{10.17}$$

则称 T 为一个**收缩算子**. 常数 α 叫作 T 的**收缩常数**.

注 不等式 (10.17) 成立的充分必要条件是: 对所有 $x \in J$, 都有 $|T(\varphi)(x) - T(\psi)(x)| \leqslant \alpha\|\varphi - \psi\|$.

例 1 令 T 为由 $T(\varphi)(x) = \lambda\varphi(x)$ 定义的算子, 此处 λ 是一个常数. 由于

$$|T(\varphi)(x) - T(\psi)(x)| = |\lambda|\,|\varphi(x) - \psi(x)|,$$

因此我们有 $\|T(\varphi) - T(\psi)\| = |\lambda|\|\varphi - \psi\|$. 因此当且仅当 $|\lambda| < 1$ 时这个算子是收缩算子, 此时可取 $|\lambda|$ 为收缩常数. □

例 2 令 $T(\varphi)(x) = b + \int_c^x f[t,\varphi(t)]\mathrm{d}t$, 此处 f 对所有 $x \in J$ 以及所有实数 y 与 z 都满足形如

$$|f(x,y) - f(x,z)| \leqslant K|y - z|$$

的 Lipschitz 条件, 其中 K 是一个正常数. 令 $L(J)$ 表示区间 J 的长度. 若 $KL(J) < 1$, 则易知 T 是以 $KL(J)$ 为收缩常数的收缩算子. 事实上, 对任意 $x \in J$, 我们都有

$$\begin{aligned}|T(\varphi)(x) - T(\psi)(x)| &= \left|\int_c^x \{f[t,\varphi(t)] - f[t,\psi(t)]\}\mathrm{d}t\right| \leqslant K\left|\int_c^x |\varphi(t) - \psi(t)|\mathrm{d}t\right| \\ &\leqslant K\|\varphi - \psi\|\left|\int_c^x \mathrm{d}t\right| \leqslant KL(J)\|\varphi - \psi\|.\end{aligned}$$

若 $KL(J) < 1$, 则 T 是以 $\alpha = KL(J)$ 为收缩常数的收缩算子. □

*10.10 关于收缩算子的不动点定理

下述定理指出任何收缩算子都有唯一的一个不动点.

定理 10.5 令 $T: C(J) \to C(J)$ 为一个收缩算子. 则在 $C(J)$ 中存在唯一的一个函数 φ 使得

$$T(\varphi) = \varphi. \tag{10.18}$$

证明 设 φ_0 为 $C(J)$ 中任意函数, 由递推公式

$$\varphi_{n+1} = T(\varphi_n), \quad n = 0, 1, 2, \cdots$$

定义函数序列 $\{\varphi_n\}$, 则对每个 n 都有 $\varphi_{n+1} \in C(J)$. 我们要证明序列 $\{\varphi_n\}$ 收敛到 $C(J)$ 中的一个极限函数 φ, 所用的方法与定理 10.3 的证明中所用的方法类似. 将各个 φ_n 表成和式,

$$\varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{n-1}\{\varphi_{k+1}(x) - \varphi_k(x)\}, \tag{10.19}$$

然后由无穷级数

$$\varphi_0(x) + \sum_{k=0}^{\infty}\{\varphi_{k+1}(x) - \varphi_k(x)\} \tag{10.20}$$

在 J 上的一致收敛性证明序列 $\{\varphi_n\}$ 收敛. 下面证明级数 (10.20) 的和就是所要求的不动点.

将无穷级数 (10.20) 与收敛的几何级数

$$M\sum_{k=0}^{\infty}\alpha^k$$

相比较即可证得它的一致收敛性, 此处 $M = \|\varphi_0\| + \|\varphi_1\|$, α 是 T 的一个收缩常数. 为此, 我们只需要证明对所有 $x \in J$ 及每个 $k \geqslant 1$, 不等式

$$|\varphi_{k+1}(x) - \varphi_k(x)| \leqslant M\alpha^k \tag{10.21}$$

成立即可. 注意到

$$|\varphi_{k+1}(x) - \varphi_k(x)| = |T(\varphi_k)(x) - T(\varphi_{k-1})(x)| \leqslant \alpha\|\varphi_k - \varphi_{k-1}\|. \tag{10.22}$$

因此若能证明对每一个 $k \geqslant 1$ 都有

$$\|\varphi_k - \varphi_{k-1}\| \leqslant M\alpha^{k-1}, \tag{10.23}$$

则不等式 (10.21) 成立. 下面用归纳法证明不等式 (10.23). 首先, 我们有

$$\|\varphi_1 - \varphi_0\| \leqslant \|\varphi_1\| + \|\varphi_0\| = M,$$

即当 $k = 1$ 时不等式 (10.23) 成立. 今设不等式 (10.23) 对 k 成立, 则由 (10.22) 得

$$|\varphi_{k+1}(x) - \varphi_k(x)| \leqslant \alpha\|\varphi_k - \varphi_{k-1}\| \leqslant M\alpha^k.$$

由于上式对 J 中的每一个 x 都成立, 因此也有

$$\|\varphi_{k+1} - \varphi_k\| \leqslant M\alpha^k,$$

即不等式 (10.23) 对 $k+1$ 也成立. 因此由归纳法即得不等式 (10.23), 从而级数 (10.20) 在 J 上一致收敛. 令 φ 表示这个级数的和, 则得

$$\varphi(x) = \lim_{n\to\infty}\varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{\infty}\{\varphi_{k+1}(x) - \varphi_k(x)\}. \tag{10.24}$$

由于 φ 是 J 上连续函数的一致收敛级数的和, 因此 φ 在 J 上连续. 为证明 φ 是 T 的一个不动点, 我们来比较 $T(\varphi)$ 与 $\varphi_{n+1} = T(\varphi_n)$. 利用 T 的收缩性质我们

得

$$|T(\varphi)(x)-\varphi_{n+1}(x)|=|T(\varphi)(x)-T(\varphi_n)(x)|\leqslant\alpha|\varphi(x)-\varphi_n(x)|. \tag{10.25}$$

但由 (10.19) 与 (10.24) 两式我们有

$$|\varphi(x)-\varphi_n(x)|=|\sum_{k=n}^{\infty}\{\varphi_{k+1}(x)-\varphi_k(x)\}|\leqslant\sum_{k=n}^{\infty}|\varphi_{k+1}(x)-\varphi_k(x)|\leqslant M\sum_{k=n}^{\infty}\alpha^k,$$

在上式的最后一步我们用了不等式 (10.21). 因此由 (10.25) 即得

$$|T(\varphi)(x)-\varphi_{n+1}(x)|\leqslant M\sum_{k=n}^{\infty}\alpha^{k+1}.$$

当 $n\to\infty$, 上式右边的级数趋于 0, 因此 $\varphi_{n+1}(x)\to T(\varphi)(x)$. 但由于当 $n\to\infty$ 时 $\varphi_{n+1}(x)\to\varphi(x)$, 因此我们证明了对每个 $x\in J$ 都有 $\varphi(x)=T(\varphi)(x)$. 于是 $\varphi=T(\varphi)$, 即 φ 是一个不动点.

最后, 我们来证明不动点 φ 的唯一性. 设 ψ 为 $C(J)$ 中另一个函数, 使得 $T(\psi)=\psi$. 则我们有

$$\|\varphi-\psi\|=\|T(\varphi)-T(\psi)\|\leqslant\alpha\|\varphi-\psi\|.$$

由此得 $(1-\alpha)\|\varphi-\psi\|\leqslant 0$, 再由 $\alpha<1$ 即得不等式 $\|\varphi-\psi\|\leqslant 0$. 于是再由 $\|\varphi-\psi\|\geqslant 0$ 即得 $\|\varphi-\psi\|=0$, 从而 $\varphi=\psi$. 由此即得不动点定理. □

*10.11 不动点定理的应用

前面已经指出, 不动点定理给出了关于线性微分方程组的一个一般的存在唯一性定理. 为说明它应用的广泛, 我们用不动点定理证明分析中的两个重要结果. 第一个结果给出了一个形如 $f(x,y)=0$ 的方程可以将 y 定义为 x 的函数的一个充分条件.

定理 10.6 (隐函数定理) *设 f 为定义在矩形带*

$$R=\{(x,y):a\leqslant x\leqslant b,-\infty<y<+\infty\}$$

上的二元函数. 假定偏导数 $D_2f(x,y)$[①] *存在且对 R 中所有的 (x,y) 都满足一个形如*

$$0<m\leqslant D_2f(x,y)\leqslant M \tag{10.26}$$

的不等式, 其中 m 与 M 为正常数, $m\leqslant M$. 再设对每一个在 $[a,b]$ 上连续的函数 φ, 复合函数 $g(x)=f[x,\varphi(x)]$ 在 $[a,b]$ 上也连续. 则存在唯一的一个在 $[a,b]$ 上连续的函数 $y=Y(x)$, 使得对所有 $x\in[a,b]$ 都有

$$f[x,Y(x)]=0. \tag{10.27}$$

---

① $D_2f(x,y)$ 表示 x 固定时, $f(x,y)$ 关于 y 的导数.

注 我们将这个结果表述为方程 $f(x,y)=0$ 在 $[a,b]$ 中定义了 x 的一个隐函数 y.

证明 令 C 为 $[a,b]$ 上的连续函数组成的线性空间, 由方程

$$T(\varphi)(x)=\varphi(x)-\frac{1}{M}f[x,\varphi(x)]$$

定义算子 $T:C\to C$, 此处 x 为 $[a,b]$ 中任意元素, M 是 (10.26) 中的正常数. 当 $\varphi\in C$ 时都有 $T(\varphi)\in C$. 我们要证明 T 是一个收缩算子. 一旦证明了这一点, 即可知 T 在 C 中有唯一的一个不动点 Y, 对这个 Y 我们有 $Y=T(Y)$, 即对每一个 $x\in[a,b]$, 都有

$$Y(x)=Y(x)-\frac{1}{M}f[x,Y(x)].$$

由此即得 (10.27).

为证明 T 是收缩算子, 我们考虑差

$$T(\varphi)(x)-T(\psi)(x)=\varphi(x)-\psi(x)-\frac{f[x,\varphi(x)]-f[x,\psi(x)]}{M}. \tag{10.28}$$

由关于导数的中值定理得

$$f[x,\varphi(x)]-f[x,\psi(x)]=\mathrm{D}_2f[x,z(x)][\varphi(x)-\psi(x)],$$

此处 $z(x)$ 落在 $\varphi(x)$ 与 $\psi(x)$ 之间. 因此由 (10.28) 式得

$$T(\varphi)(x)-T(\psi)(x)=[\varphi(x)-\psi(x)]\left(1-\frac{\mathrm{D}_2f[x,z(x)]}{M}\right). \tag{10.29}$$

由条件 (10.26) 得

$$0\leqslant\left(1-\frac{\mathrm{D}_2f[x,z(x)]}{M}\right)\leqslant 1-\frac{m}{M}$$

从而由等式 (10.29) 得不等式

$$|T(\varphi)(x)-T(\psi)(x)|\leqslant|\varphi(x)-\psi(x)|\left(1-\frac{m}{M}\right)\leqslant\alpha\|\varphi-\psi\|, \tag{10.30}$$

其中 $\alpha=1-\frac{m}{M}$. 由于 $0<m\leqslant M$, 因此我们有 $0\leqslant\alpha<1$. 不等式 (10.30) 对任意 $x\in[a,b]$ 都成立. 因此 T 是一个收缩算子. 证明完成. □

不动点定理的第二个应用是建立积分方程

$$\varphi(x)=\psi(x)+\lambda\int_a^b K(x,t)\varphi(t)\mathrm{d}t \tag{10.31}$$

的一个存在性定理. 此处 ψ 是一个给定的在 $[a,b]$ 上连续的函数, λ 是给定常数, K 是一个给定的在正方形

$$S=\{(x,y):a\leqslant x\leqslant b,a\leqslant y\leqslant b\}$$

上有界的函数. K 叫作此积分方程的**核**(kernel). 令 C 为由 $[a,b]$ 上全体连续函数组成的线性空间. 假定核 K 使得由

$$T(\varphi)(x)=\psi(x)+\lambda\int_a^b K(x,t)\varphi(t)\mathrm{d}t$$

给出的算子 T 将 C 映入 C. 换言之, 对所有 $\varphi \in C$ 都有 $T(\varphi) \in C$. C 中任一满足方程 (10.31) 的函数 φ 就叫作积分方程的一个解.

定理 10.7 (积分方程的存在唯一性定理) *若在上述条件下, 对所有 $(x,y) \in S$, 都有*

$$|K(x,y)| \leqslant M, \tag{10.32}$$

则对所有满足条件

$$|\lambda| < \frac{1}{M(b-a)} \tag{10.33}$$

的 λ, 都存在 C 中唯一的一个函数 φ 满足积分方程 (10.31).

证明 我们来证明 T 是一个收缩算子. 取 C 中任意两个函数 φ_1 与 φ_2, 考虑差

$$T(\varphi_1)(x) - T(\varphi_2)(x) = \lambda \int_a^b K(x,t)[\varphi_1(t) - \varphi_2(t)]\mathrm{d}t.$$

利用 (10.32) 中给出的核 $K(x,t)$ 的上界, 我们有

$$|T(\varphi_1)(x) - T(\varphi_2)(x)| = |\lambda| M(b-a)\|\varphi_1 - \varphi_2\| = \alpha\|\varphi_1 - \varphi_2\|,$$

其中 $\alpha = |\lambda| M(b-a)$. 由不等式 (10.33) 得 $0 \leqslant \alpha < 1$, 因此 T 是一个收缩算子, 收缩常数是 α. 从而 T 在 C 中有唯一的一个不动点 φ, 这个函数 φ 满足积分方程 (10.31). □

习 题 解 答

0.8 习题

3. (b) 由于 $(a,b)(a,b)=(a^2,b^2)$, 因此一个复数的平方绝不可能等于 $(-1,0)$. 两个都不等于零的复数的乘积有可能为零. 例如 $(1,0)(0,1)=(0,0)$. 因此, 通过先将 $f(x)$ 分解成线性因子的乘积再令各个线性因子等于零来解代数方程 $f(x)=0$ 的通常方法对这样定义的乘法不再适用.

0.11 习题

1. (a) $2i$ (b) $-i$ (c) $\frac{1}{2}-\frac{1}{2}i$ (d) $18+i$ (e) $-\frac{1}{5}+\frac{3}{5}i$ (f) $1+i$ (g) 0 (h) $1+i$
2. (a) $\sqrt{2}$ (b) 5 (c) 1 (d) 1 (e) $\sqrt{2}$ (f) $\sqrt{65}$
3. (a) $r=2,\theta=\frac{1}{2}\pi$ (b) $r=3,\theta=-\frac{1}{2}\pi$ (c) $r=1,\theta=\pi$ (d) $r=1,\theta=0$
   (e) $r=2\sqrt{3},\theta=5\pi/6$ (f) $r=1,\theta=\frac{1}{4}\pi$ (g) $r=2\sqrt{2},\theta=\frac{1}{4}\pi$ (h) $r=2\sqrt{2},\theta=-\frac{1}{4}\pi$
   (i) $r=\frac{1}{2}\sqrt{2},\theta=-\frac{1}{4}\pi$ (j) $r=\frac{1}{2},\theta=-\frac{1}{2}\pi$
4. (a) $y=0$, x 任意 (b) $x\geqslant 0$, $y=0$ (c) 所有 x 和 y (d) $x=0$, y 任意; 或 $y=0$, x 任意
   (e) $x=1$, $y=0$ (f) $x=1$, $y=0$

0.26 习题

1. $(1,\frac{7}{3})$ 与 $(3,1)$.
2. $a=-2,b=4$.
3. $a=1;b=0,c=-1$.
4. (b) $f'(t)=144-32t$.
   (c) 在运动的中点, 即 $t=9/2$ 时, 速度为 0.
   (d) 加速度是 $f''(t)=-32$.
5. (a) $f'(x)=\frac{1-\ln x}{x^2}>0$, 若 $0<x<e$; $f''(x)=\frac{-3-2\ln x}{x^3}>0$, 若 $x>e^{3/2}$.
   (b) $\frac{1}{2}(\ln x)^2$ 的导数是 $\frac{\ln x}{x}$, 因此, $\int_1^t \frac{\ln x}{x}dx=\frac{1}{2}(\ln t)^2$; 常数是 $\frac{1}{2}$.
   (c) $m=\frac{1}{2e}$, 切点是 $(\sqrt{e},\frac{1}{2e})$.
   (d) $b=\sqrt{e}$.
6. $p(x)=3+x+\frac{1}{6}x^3$
8. (b) $a=1,b=0$.

1.4 习题

1. (a) (5, 0, 9) (b) $(-3,6,3)$ (c) $(3,-1,4)$ (d) $(-7,24,21)$ (e) (0,0,0)
5. $x=\frac{1}{5}(3c_1-c_2),y=\frac{1}{5}(2c_2-c_1)$
6. (a) $(x+z,x+y+z,x+y)$ (c) $x=2,y=1,z=-1$

7. (a) $(x+2z, x+y+z, x+y+z)$ (b) 一个例子：$x=-2$, $y=z=1$
8. (a) $(x+z, x+y+z, x+y, y)$ (c) $x=-1, y=4, z=2$
12. 平行四边形的对角线互相平分.

1.8 习题

1. (a) -6 (b) 2 (c) 6 (d) 0 (e) 4
2. (a) $(\boldsymbol{A}\cdot\boldsymbol{B})\boldsymbol{C}=(21,28,-35)$ (b) $\boldsymbol{A}\cdot(\boldsymbol{B}+\boldsymbol{C})=64$ (c) $(\boldsymbol{A}+\boldsymbol{B})\cdot\boldsymbol{C}=72$

   (d) $\boldsymbol{A}(\boldsymbol{B}\cdot\boldsymbol{C})=(30,60,-105)$ (e) $\boldsymbol{A}/(\boldsymbol{B}\cdot\boldsymbol{C})=(\frac{2}{15},\frac{4}{15},\frac{-7}{15})$
5. 一个例子：$(1,-5,-3)$
6. 一个例子：$x=-2$, $y=1$
7. $\boldsymbol{C}=\frac{4}{9}(-1,-2,2)$, $\boldsymbol{D}=\frac{1}{9}(22,-1,10)$
8. $\boldsymbol{C}=\frac{1}{11}(1,2,3,4,5)$, $\boldsymbol{D}=(\frac{5}{11},\frac{7}{44},\frac{1}{33},\frac{-5}{88},\frac{-7}{55})$
9. (a) $\sqrt{74}$ (b) $\sqrt{14}$ (c) $\sqrt{53}$ (d) 5
10. (a) $(1,-1)$ 或 $(-1,1)$ (b) $(1,1)$ 或 $(-1,-1)$ (c) $(3,2)$ 或 $(-3,-2)$ (d) $(b,-a)$ 或 $(-b,a)$
11. (a) $\frac{1}{\sqrt{42}}(4,-1,5)$ (b) $\frac{1}{\sqrt{14}}(-2,-3,1)$ (c) $\frac{1}{\sqrt{2}}(1,0,1)$

    (d) $\frac{1}{\sqrt{42}}(-5,-4,-1)$ (e) $\frac{1}{\sqrt{42}}(-1,-5,4)$
12. $\boldsymbol{A}$ 与 $\boldsymbol{B}$, $\boldsymbol{C}$ 与 $\boldsymbol{D}$, $\boldsymbol{C}$ 与 $\boldsymbol{E}$, $\boldsymbol{D}$ 与 $\boldsymbol{E}$
13. (a) $(2,-1)$ 与 $(-2,1)$ (b) $(2,1)$ 与 $(-2,-1)$ (c) $(1,2)$ 与 $(-1,-2)$

    (d) $(1,2)$ 与 $(-1,-2)$
14. 一个例子：$\boldsymbol{C}=(8,1,1)$
15. 一个例子：$\boldsymbol{C}=(1,-5,-3)$
16. $P=\frac{11}{25}(3,4)$, $Q=\frac{2}{25}(-4,3)$
17. $P=\frac{5}{2}(1,1,1,1)$, $Q=\frac{1}{2}(-3,-1,1,3)$
18. $\pm\frac{1}{\sqrt{2}}(0,1,1)$
20. 任意一个平行四边形的边长的平方和等于两条对角线长的平方和.
22. $4; 12\sqrt{2}$
23. $\boldsymbol{C}=\frac{1}{11}(1,2,3,4,5)$, $\boldsymbol{D}=\frac{1}{11}(10,\frac{7}{2},\frac{2}{3},\frac{-5}{4},\frac{-14}{5})$
24. $\boldsymbol{C}=t\boldsymbol{A}$, $\boldsymbol{D}=\boldsymbol{B}-t\boldsymbol{A}$, 此处 $t=(\boldsymbol{A}\cdot\boldsymbol{B})/(\boldsymbol{A}\cdot\boldsymbol{A})$

1.11 习题

1. $\frac{11}{9}\boldsymbol{B}$
2. $\frac{5}{2}\boldsymbol{B}$
3. (a) $\frac{6}{7},\frac{3}{7},\frac{-2}{7}$ (b) $(\frac{6}{7},\frac{3}{7},\frac{-2}{7})$ 和 $(\frac{-6}{7},\frac{-3}{7},\frac{2}{7})$
5. $0,\sqrt{\frac{35}{41}},\sqrt{\frac{6}{41}}$
6. $7\pi/8$
8. $\pi/6$
9. 0
10. (b) 若 $\cos\theta=1$, 则等式对所有 x 与 y 都成立; 否则 $x=y=0$ 是仅有的解.

14. 除 (b) 之外的全部
17. (c) 除了定理 1.4(a) 之外的全部
18. (a) 全部

1.15 习题

1. (a) $x=y=\frac{1}{2}$ (b) $x=-\frac{1}{2}, y=\frac{1}{2}$ (c) $x=4, y=-1$ (d) $x=1, y=6$
2. $x=\frac{1}{4}, y=\frac{7}{8}$
3. $x=3, y=-4$
7. 所有 $t\neq 0$
9. (c) $7i-4(i+j)$
10. (b) $j=\boldsymbol{B}-\boldsymbol{A}, k=\frac{1}{3}(\boldsymbol{C}-\boldsymbol{B})$ (c) $\frac{1}{3}(15\boldsymbol{A}-14\boldsymbol{B}+5\boldsymbol{C})$
11. $\{\boldsymbol{A}\}, \{\boldsymbol{B}\}, \{\boldsymbol{C}\}, \{\boldsymbol{D}\}, \{\boldsymbol{A},\boldsymbol{B}\}, \{\boldsymbol{A},\boldsymbol{C}\}, \{\boldsymbol{A},\boldsymbol{D}\}, \{\boldsymbol{B},\boldsymbol{C}\}, \{\boldsymbol{C},\boldsymbol{D}\}$
12. (a) 线性无关 (b) 一个例子：$\boldsymbol{D}=\boldsymbol{A}$ (c) 一个例子：$\boldsymbol{E}=(0,0,0,1)$ (d) 对选择 $\boldsymbol{E}=(0,0,0,1)$ 我们有 $\boldsymbol{X}=2\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{C}+3\boldsymbol{E}$
13. (c) $t=0, \sqrt{2}, -\sqrt{2}$
14. (a) $\{(1,0,1,0),(0,1,0,1),(2,0,-1,0)\}$ (b) 给定的集合 (c) 给定的集合
17. $\{(0,1,1),(1,1,1),(0,1,0)\}, \{(0,1,1),(1,1,1),(0,0,1)\}$
18. $\{(1,1,1,1),(0,1,1,1),(0,0,1,1),(0,0,0,1)\}, \{(1,1,1,1),(0,1,1,1),(0,1,0,0),(0,0,1,0)\}$
19. $L(U)\subseteq L(T)\subseteq L(S)$
20. 一个例子：$A=\{I_1,\cdots,I_n\}$, $B=\{I_1+I_2, I_2+I_3,\cdots,I_{n-1}+I_n, I_n+I_1\}$

1.17 习题

1. (a) $-1-i$ (b) $-1+i$ (c) $1-i$ (d) $-1+i$ (e) $-1-i$ (f) $2-i$ (g) $-i$ (h) $-1+2i$ (i) $-3-2i$ (j) $2i$
2. 一个例子：$(1+i,-5-3i,1-3i)$
8. $\pi/3$
9. $3\boldsymbol{A}-\boldsymbol{B}+2\boldsymbol{C}$

2.6 习题

1. (b), (d) 与 (e)
2. (a) 与 (e)
3. (c), (d) 与 (e)
4. (b), (e) 与 (f)
5. (a) 不共线 (b) 不共线 (c) 不共线
6. $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{F}$ 共线
7. 相交于 $(5,9,13)$
8. (b) 不相交
9. (a) $9t^2+8t+9$ (b) $\frac{1}{3}\sqrt{65}$

2.9 习题

1. (c) 与 (e)
2. (a), (b) 与 (c)

3. (a) $x=1+t,\ y=2+s+t,\ z=1+4t$ (b) $x=s+t,\ y=1+s,\ z=s+4t$
4. (a) $(1,2,0)$ 与 $(2,-3,-3)$ (b) $M=\{(1,2,0)+s(1,1,2)+t(-2,4,1)\}$
6. (a), (b) 与 (c) $x-2y+z=-3$
7. (a) $(0,-2,-1)$ 与 $(-1,-2,2)$ (b) $M=\{(0,-2,-1)+s(-1,0,3)+t(3,3,6)\}$
8. 两个例子：$(-5,2,6)$ 与 $(-14,3,17)$
9. (a) 平行 (b) 两个例子：$(1,0,-1)$ 与 $(-1,0,1)$
10. $(-2,\frac{5}{2},-\frac{7}{2})$
11. (a), (b) 与 (c) 不平行
13. $x-y=-1$

2.12 习题

1. (a) $(-2,3,-1)$ (b) $(4,-5,3)$ (c) $(4,-4,2)$ (d) $(8,10,4)$ (e) $(8,3,-7)$ (f) $(10,11,5)$ (g) $(-2,-8,-12)$ (h) $(2,-2,0)$ (i) $(-2,0,4)$
2. (a) $\pm\frac{1}{\sqrt{26}}(-4,3,1)$ (b) $\pm\frac{1}{\sqrt{2054}}(-41,-18,7)$ (c) $\pm\frac{1}{\sqrt{6}}(1,2,1)$
3. (a) $\frac{15}{2}$ (b) $\frac{3}{2}\sqrt{35}$ (c) $\frac{1}{2}\sqrt{3}$
4. $8\boldsymbol{i}+\boldsymbol{j}-2\boldsymbol{k}$
6. (b) $\cos\theta$ 为负 (c) $\sqrt{5}$
9. (a) 有一个解是 $\boldsymbol{B}=-\boldsymbol{i}-3\boldsymbol{k}$ (b) $\boldsymbol{i}-\boldsymbol{j}-\boldsymbol{k}$ 是仅有的解
11. (a) 三种可能：$\boldsymbol{D}=\boldsymbol{B}+\boldsymbol{C}-\boldsymbol{A}=(0,0,2)$, $\boldsymbol{D}=\boldsymbol{A}+\boldsymbol{C}-\boldsymbol{B}=(4,-2,2)$, $\boldsymbol{D}=\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{C}=(-2,2,0)$ (b) $\frac{1}{2}\sqrt{6}$
12. $-4;\ 8\sqrt{3};\ -\frac{1}{2}\sqrt{3}$

2.15 习题

1. (a) 96 (b) 27 (c) -84
2. $0,\sqrt{2},-\sqrt{2}$
3. 2
6. (a) $(2b-1)\boldsymbol{i}+b\boldsymbol{j}+c\boldsymbol{k}$, 此处 b 与 c 是任意实数 (b) $-\frac{1}{5}\boldsymbol{i}+\frac{2}{5}\boldsymbol{j}$
11. $-3i+2j+5k$
14. (b) $1/3$
15. (b) $\sqrt{2005/41}$
17. $x=1, y=-1, z=2$
18. $x=1, y=-1, z=2$
19. $x=1, y=4, z=1$

2.18 习题

1. (a) $(-7,2,-2)$ (b) $-7x+2y-2z=0$ (c) $-7x+2y-2z=-9$
2. (a) $(\frac{1}{3},\frac{2}{3},-\frac{2}{3})$ (b) $-7,-\frac{7}{2},\frac{7}{2}$ (c) $\frac{7}{3}$ (d) $(-\frac{7}{9},-\frac{14}{9},\frac{14}{9})$
3. $3x-y+2z=-5; 9/\sqrt{14}$
4. (b) $\frac{19}{18}\sqrt{6}$
5. (a) $(1,2,-2)$ (b) $x+2y-2z=5$ (c) $\frac{5}{3}$
6. $10x-3y-7z+17=0$
7. 两个角：$\pi/3$ 与 $2\pi/3$

8. $x + 2y + 9z + 55 = 0$
9. $X(t) = (2, 1, -3) + t(4, -3, 1)$
10. (b) $N = (1, 3, -2)$ (c) $t = 1$ (d) $2x + 3y + 2z + 15 = 0$ (e) $x + 3y - 2z + 19 = 0$
11. $x + \sqrt{2}y + z = 2 + \sqrt{2}$
12. 6
13. $\frac{1}{\sqrt{122}}(7, -8, -3)$
14. $x - y + z = 2$
15. $(\frac{3}{2}, 0, \frac{1}{2})$
17. $X(t) = (1, 2, 3) + t(1, -2, 1)$
19. (b) $P = -\frac{1}{25}(5, -14, 2)$
20. $2x - y + 2z - 8 = 0$
21. (b) 1
23. $3x - z = 0$
24. $2x - 3z = 2$

2.22 习题

3. $r = ed/(1 - e\sin\theta); r = -ed/(1 + e\sin\theta)$
4. $e = 1$, $d = 2$
5. $e = \frac{1}{2}$, $d = 6$
6. $e = \frac{1}{3}$, $d = 6$
7. $e = 2$, $d = 1$
8. $e = 2$, $d = 2$
9. $e = 1$, $d = 4$
10. $d = 5$, $r = 25/(10 + 3\cos\theta + 4\sin\theta)$
11. $d = 5$, $r = 25/(5 + 4\cos\theta + 3\sin\theta)$
12. $d = \frac{1}{2}\sqrt{2}$, $r = 1/(\cos\theta + \sin\theta + \frac{1}{2}\sqrt{2})$, $r = 1/(\cos\theta + \sin\theta - \frac{1}{2}\sqrt{2})$
13. (a) $r = 1.5 \times 10^8/(1 + \cos\theta)$; 7.5×10^7 英里
    (b) $r = 5 \times 10^7/(1 - \cos\theta)$; 2.5×10^7 英里

2.27 习题

1. 中心坐标为 $(0, 0)$; 焦点坐标为 $(\pm 8, 0)$; 顶点坐标为 $(\pm 10, 0)$; $e = \frac{4}{5}$
2. 中心坐标为 $(0, 0)$; 焦点坐标为 $(0, \pm 8)$; 顶点坐标为 $(0, \pm 10)$; $e = \frac{4}{5}$
3. 中心坐标为 $(2, -3)$; 焦点坐标为 $(2 \pm \sqrt{7}, -3)$; 顶点坐标为 $(6, -3)$, $(-2, -3)$; $e = \frac{\sqrt{7}}{4}$
4. 中心坐标为 $(0, 0)$; 焦点坐标为 $(\pm\frac{4}{3}, 0)$; 顶点坐标为 $(\pm\frac{5}{3}, 0)$; $e = \frac{4}{5}$
5. 中心坐标为 $(0, 0)$; 焦点坐标为 $(\pm\sqrt{3}/6, 0)$; 顶点坐标为 $(\pm\sqrt{3}/3, 0)$; $e = \frac{1}{2}$
6. 中心坐标为 $(-1, -2)$; 焦点坐标为 $(-1, 1)$, $(-1, -5)$; 顶点坐标为 $(-1, 3)$, $(-1, -7)$; $e = \frac{3}{5}$
7. $7x^2 + 16y^2 = 7$
8. $\frac{(x+3)^2}{16} + \frac{(y-4)^2}{9} = 1$
9. $\frac{(x+3)^2}{9} + \frac{(y-4)^2}{16} = 1$
10. $\frac{(x+4)^2}{9} + \frac{(y-2)^2}{1} = 1$

11. $\frac{(x-8)^2}{25}+\frac{(y+2)^2}{9}=1$

12. $\frac{(x-2)^2}{16}+\frac{(y-1)^2}{4}=1$

13. 中心坐标为 $(0,0)$; 焦点坐标为 $(\pm2\sqrt{41},0)$; 顶点坐标为 $(\pm10,0)$; $e=\frac{\sqrt{41}}{5}$

14. 中心坐标为 $(0,0)$; 焦点坐标为 $(0,\pm2\sqrt{41})$; 顶点坐标为 $(0,\pm10)$; $e=\frac{\sqrt{41}}{5}$

15. 中心坐标为 $(-3,3)$; 焦点坐标为 $(-3\pm\sqrt{5},3)$; 顶点坐标为 $(-1,3)$, $(-5,3)$; $e=\frac{\sqrt{5}}{2}$

16. 中心坐标为 $(0,0)$; 焦点坐标为 $(\pm5,0)$; 顶点坐标为 $(\pm4,0)$; $e=\frac{5}{4}$

17. 中心坐标为 $(0,0)$; 焦点坐标为 $(0,\pm3)$; 顶点坐标为 $(0,\pm2)$; $e=\frac{3}{2}$

18. 中心坐标为 $(1,-2)$; 焦点坐标为 $(1\pm\sqrt{13},-2)$; 顶点坐标为 $(3,-2)$, $(-1,-2)$; $e=\frac{\sqrt{13}}{2}$

19. $\frac{x^2}{4}-\frac{y^2}{12}=1$

20. $y^2-x^2=1$

21. $\frac{x^2}{4}-\frac{y^2}{16}=1$

22. $\frac{(y-4)^2}{1}-\frac{(x+1)^2}{3}=1$

23. $\frac{8(y+3)^2}{27}-\frac{5(x-2)^2}{27}=1$

24. $4x^2-y^2=11$

25. 顶点坐标为 $(0,0)$; 准线方程为 $x=2$; 对称轴方程为 $y=0$

26. 顶点坐标为 $(0,0)$; 准线方程为 $x=-\frac{3}{4}$; 对称轴方程为 $y=0$

27. 顶点坐标为 $(\frac{1}{2},1)$; 准线方程为 $x=-\frac{5}{2}$; 对称轴方程为 $y=1$

28. 顶点坐标为 $(0,0)$; 准线方程为 $y=-\frac{3}{2}$; 对称轴方程为 $x=0$

29. 顶点坐标为 $(0,0)$; 准线方程为 $y=2$; 对称轴方程为 $x=0$

30. 顶点坐标为 $(-2,-\frac{9}{4})$; 准线方程为 $y=-\frac{13}{4}$; 对称轴方程为 $x=-2$

31. $x^2=-y$

32. $y^2=8x$

33. $(x+4)^2=-8(y-3)$

34. $(y+1)^2=5(x-\frac{7}{4})$

35. $(x-\frac{3}{2})^2=2(y+\frac{1}{8})$

36. $(y-3)^2=-8(x-1)$

37. $x^2-4xy+4y^2+40x+20y-100=0$

38. $x^2-2xy+y^2-2x-2y=1$

39. $\frac{x^2}{12}+\frac{y^2}{16}=1$

40. $y^2-4x^2-4y+4x=0$

2.28 关于二次曲线的综合性习题

1. (a) $e=\sqrt{2/(p+2)}$; 焦点坐标为 $(\sqrt{2},0)$ 和 $(-\sqrt{2},0)$ (b) $6x^2-3y^2=4$

5. (b) $y=Cx^2, C\neq0$

6. $(x-\frac{2}{5})^2+(y-\frac{4}{5})^2=\frac{4}{5}$

11. $\pm\sqrt{\frac{23}{3}}$

12. $\frac{2}{3}bh$

13. 16π

14. (a) $\frac{8}{3}$ (b) 2π (c) $48\pi/5$

16. $A=\frac{1}{2}(1+\sqrt{5})B$

17. $(4,8)$

3.5 习题

1.	是	2.	不是	3.	是	4.	不是	5.	是
6.	是	7.	是	8.	是	9.	是	10.	是
11.	是	12.	是	13.	不是	14.	是	15.	是
16.	是	17.	是	18.	不是	19.	是	20.	是

23. (a) 不是 (b) 不是 (c) 不是 (d) 不是

25. (a) 是 (b) 不是 (c) 是 (d) 是

3.10 习题

1.	是;2 维	2.	是;2 维	3.	是;2 维	4.	是; 2 维	5.	是; 1 维
6.	不是	7.	不是	8.	不是	9.	是; 1 维	10.	是; 1 维
11.	是; n 维	12.	是; n 维	13.	是; n 维				

14. 是; 若 n 为偶数则 $\dim=1+\frac{1}{2}n$, 若 n 为奇数则 $\dim=\frac{1}{2}(n+1)$
15. 是; 若 n 为偶数则 $\dim=\frac{1}{2}n$, 若 n 为奇数则 $\dim=\frac{1}{2}(n+1)$
16. 不是
17. 是; $k+1$ 维
18. 不是
19. 是; n 维
20. 是; $n-1$ 维
21. 是; n 维
22. 是; n 维
23. (a) $\dim=3$ (b) $\dim=3$ (c) $\dim=2$ (d) $\dim=2$
24. (a) 若 $a\neq 0$ 且 $b\neq 0$, 则集合线性无关; $\dim=3$; 若 a 和 b 中有一个为零, 则集合线性相关; $\dim=2$.
(b) 线性无关, $\dim=2$ (c) 若 $a\neq 0$, 则线性无关, $\dim=3$; 若 $a=0$, 则线性相关, $\dim=2$
(d) 线性无关; $\dim=3$ (e) 线性相关; $\dim=2$ (f) 线性无关; $\dim=2$ (g) 线性无关; $\dim=2$
(h) 线性相关; $\dim=2$ (i) 线性无关; $\dim=2$ (j) 线性无关; $\dim=2$

3.13 习题

1. (a) 不是内积 (b) 不是内积 (c) 不是内积 (d) 不是内积 (e) 是内积

10. (b) $\frac{(n+1)(2n+1)}{6n}a+\frac{n+1}{2}b$ (c) $g(t)=a(t-\frac{2n+1}{3n})$, a 任意

11. (a) $\frac{1}{2}\sqrt{e^2+1}$ (b) $g(x)=b(x-\frac{1}{4}(e^2+1))$, b 任意

13. (c) 43 (d) $g(t)=a(1-\frac{2}{3}t)$, a 任意

14. (a) 不是内积 (b) 不是内积 (c) 不是内积 (d) 不是内积

15. (c) $n!/2^{n+1}$

16. (c) 1　(d) $e^{\frac{1}{2}}-1$

3.17 习题

1. (a) 与 (b) $\frac{1}{3}\sqrt{3}(1,1,1)$, $\frac{1}{6}\sqrt{6}(1,-2,1)$
2. (a) $\frac{1}{2}\sqrt{2}(1,1,0,0)$, $\frac{1}{6}\sqrt{6}(-1,1,2,0)$, $\frac{1}{6}\sqrt{3}(1,-1,1,3)$

   (b) $\frac{1}{3}\sqrt{3}(1,1,0,1)$, $\frac{1}{\sqrt{42}}(1,-2,6,1)$
6. $\frac{2}{3}-\frac{1}{2}\ln^2 3$
7. e^2-1
8. $\frac{1}{2}(e-e^{-1})+\frac{3}{e}x; 1-7e^{-2}$
9. $\pi-2\sin x$
10. $\frac{3}{4}-\frac{1}{4}x$

4.4 习题

1. 线性; 零化度为 0, 秩为 2
2. 线性; 零化度为 0, 秩为 2
3. 线性; 零化度为 1, 秩为 1
4. 线性; 零化度为 1, 秩为 1
5. 非线性
6. 非线性
7. 非线性
8. 非线性
9. 线性; 零化度为 0, 秩为 2
10. 线性; 零化度为 0, 秩为 2
11. 线性; 零化度为 0, 秩为 2
12. 线性; 零化度为 0, 秩为 2
13. 非线性
14. 线性; 零化度为 0, 秩为 2
15. 非线性
16. 线性; 零化度为 0, 秩为 3
17. 线性; 零化度为 1, 秩为 2
18. 线性; 零化度为 1, 秩为 3
19. 非线性
20. 非线性
21. 非线性
22. 非线性
23. 线性; 零化度为 1, 秩为 2
25. 线性, 零化度为 0, 秩为 $n+1$
26. 线性; 零化度为 1, 秩为无穷大
27. 线性; 零化度为 2, 秩为无穷大
28. 线性; 零化度为无穷大, 秩为 2
29. $N(T)$ 是全体常数序列的集合; $T(V)$ 是全体极限为 0 的序列的集合
30. (d) $\{1,\cos x,\sin x\}$ 是 $T(V)$ 的一组基; $\dim T(V)=3$

   (e) $N(T)=S$

   (f) 若 $T(f)=cf$ 且 $c\neq 0$, 则 $cf\in T(V)$, 因此 $f\in T(V)$, 从而我们有 $f(x)=c_1+c_2\cos x+c_3\sin x$. 由 $T(f)=cf$ 得

   $$2\pi c_1+\pi c_2\cos x+\pi c_3\sin x=c(c_1+c_2\cos x+c_3\sin x).$$

   若 $c_1=0$, 则 $c=\pi$ 且 $f(x)=c_2\cos x+c_3\sin x$, 其中 c_2 与 c_3 是不全为零的任意实数; 若 $c_1\neq 0$, 则 $c=2\pi$ 且 $f(x)=a$, 此处 a 为不等于零的任意实数.

4.8 习题

3. 是; $x=v$, $y=u$
4. 是; $x=u$, $y=-v$
5. 不是
6. 不是

7. 不是
8. 是; $x = \ln u$, $y = \ln v$
9. 不是
10. 是; $x = u - 1$, $y = v - 1$
11. 是; $x = \frac{1}{2}(v+u)$, $y = \frac{1}{2}(v-u)$
12. 是; $x = \frac{1}{3}(v+u)$, $y = \frac{1}{3}(2v-u)$
13. 是; $x = w$, $y = v$, $z = u$
14. 不是
15. 是; $x = u$, $y = \frac{1}{2}v$, $z = \frac{1}{3}w$
16. 是; $x = u$, $y = v$, $z = w - u - v$
17. 是; $x = u - 1$, $y = v - 1$, $z = w + 1$
18. 是; $x = u - 1$, $y = v - 2$, $z = w - 3$
19. 是; $x = u$, $y = v - u$, $z = w - v$
20. 是; $x = \frac{1}{2}(u - v + w)$, $y = \frac{1}{2}(v - w + u)$, $z = \frac{1}{2}(w - u + v)$
25. $(S+T)^2 = S^2 + ST + TS + T^2$;
$(S+T)^3 = S^3 + TS^2 + STS + S^2T + ST^2 + TST + T^2S + T^3$
27. (a) $(ST)(x,y,z) = (x+y+z, x+y, x)$; $(TS)(x,y,z) = (z, z+y, z+y+x)$;
$(ST-TS)(x,y,z) = (x+y, x-z, -y-z)$; $S^2(x,y,z) = (x,y,z)$;
$T^2(x,y,z) = (x, 2x+y, 3x+2y+z)$;
$(ST)^2(x,y,z) = (3x+2y+z, 2x+2y+z, x+y+z)$;
$(TS)^2(x,y,z) = (x+y+z, x+2y+2z, x+2y+3z)$;
$(ST-TS)^2(x,y,z) = (2x+y-z, x+2y+z, -x+y+2z)$;
(b) $S^{-1}(u,v,w) = (w,v,u)$; $T^{-1}(u,v,w) = (u, v-u, w-v)$;
$(ST)^{-1}(u,v,w) = (w, v-w, u-v)$; $(TS)^{-1}(u,v,w) = (w-v, v-u, u)$;
(c) $(T-I)(x,y,z) = (0, x, x+y)$; $(T-I)^2(x,y,z) = (0,0,x)$;
$(T-I)^n(x,y,z) = (0,0,0)$ if $n \geqslant 3$
28. (a) $Rp(x) = 2$; $Sp(x) = 3 - x + x^2$; $Tp(x) = 2x + 3x^2 - x^3 + x^4$; $(ST)p(x) = 2 + 3x - x^2 + x^3$;
$(TS)p(x) = 3x - x^2 + x^3$; $(TS)^2p(x) = 3x - x^2 + x^3$; $(T^2S^2)p(x) = -x^2 + x^3$;
$(S^2T^2)p(x) = 2 + 3x - x^2 + x^3$; $(TRS)p(x) = 3x$; $(RST)p(x) = 2$
(b) $N(R) = \{p : p(0) = 0\}$; $R(V) = \{p : p$是常数$\}$; $N(S) = \{p : p$是常数$\}$; $S(V) = V$;
$N(T) = \{O\}$; $T(V) = \{p : p(0) = 0\}$
(c) $T^{-1} = S$　　(d) $(TS)^n = I - R$; $S^nT^n = I$
29. (a) $Dp(x) = 3 - 2x + 12x^2$; $Tp(x) = 3x - 2x^2 + 12x^3$;
$(DT)p(x) = 3 - 4x + 36x^2$; $(TD)p(x) = -2x + 24x^2$; $(DT-TD)p(x) = 3 - 2x + 12x^2$;
$(T^2D^2 - D^2T^2)p(x) = 8 - 192x$
(b) $p(x) = ax$, 此处 a 是任意纯量
(c) $p(x) = ax^2 + b$, 此处 a 与 b 为任意纯量
(d) V 中所有 p
32. 由于 T 将所有的常数序列映为同一个序列, 因此 T 不是一一的.

4.12 习题

1. (a) 单位矩阵 $I = (\delta_{jk})$, 此处当 $j = k$ 时 $\delta_{jk} = 1$, 当 $j \neq k$ 时 $\delta_{jk} = 0$
(b) 零矩阵 O, 它的每一个元素都是 0

(c) 矩阵 $(c\delta_{jk})$, 此处 (δ_{jk}) 是 (a) 中的单位矩阵

2. (a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

3. (a) $-5\boldsymbol{i}+7\boldsymbol{j}$, $9\boldsymbol{i}-12\boldsymbol{j}$

(b) $\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} -\frac{7}{4} & -\frac{1}{4} \\ \frac{1}{4} & \frac{7}{4} \end{bmatrix}$, $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$

4. $\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$, $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$

5. (a) $3\boldsymbol{i}+4\boldsymbol{j}+4\boldsymbol{k}$; 零化度为 0, 秩为 3 (b) $\begin{bmatrix} -1 & -1 & 2 \\ 1 & -3 & 3 \\ -1 & -5 & 5 \end{bmatrix}$

6. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & -1 & 1 \\ 2 & 1 & 0 \end{bmatrix}$

7. (a) $T(4\boldsymbol{i}-\boldsymbol{j}+\boldsymbol{k})=(0,-2)$; 零化度为 1, 秩为 2 (b) $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}$ (d) $e_1=\boldsymbol{j}$, $e_2=\boldsymbol{k}$, $e_3=\boldsymbol{i}$, $w_1=(1,1)$, $w_2=(1,-1)$

8. (a) $(5,0,-1)$; 零化度为 0, 秩为 2 (b) $\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$

(c) $e_1=\boldsymbol{i}$, $e_2=\boldsymbol{i}+\boldsymbol{j}$, $w_1=(1,0,1)$, $w_2=(0,0,2)$, $w_3=(0,1,0)$

9. (a) $(-1,-3,-1)$; 零化度为 0, 秩为 2 (b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$

(c) $e_1=\boldsymbol{i}$, $e_2=\boldsymbol{j}-\boldsymbol{i}$, $w_1=(1,0,1)$, $w_2=(0,1,0)$, $w_3=(0,0,1)$

10. (a) e_1-e_2; 零化度为 0, 秩为 2 (b) $\begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$ (c) $a=5$, $b=4$

11. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

12. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

16. $\begin{bmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 & 0 & -2 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

17. $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$, $\begin{bmatrix} -5 & -12 \\ 12 & -5 \end{bmatrix}$

19. (a) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$ (f) $\begin{bmatrix} 0 & 0 & -8 & 0 \\ 0 & 0 & 0 & -48 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

20. 取 $(x^3, x^2, x, 1)$ 为 V 的一组基, (x^2, x) 为 W 的一组基, 则 TD 的矩阵为 $\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$.

4.16 习题

1. $B+C=\begin{bmatrix} 3 & 4 \\ 0 & 2 \\ 6 & -5 \end{bmatrix}$, $AB=\begin{bmatrix} 15 & -14 \\ -15 & 14 \end{bmatrix}$, $BA=\begin{bmatrix} -1 & 4 & -2 \\ -4 & 16 & -8 \\ 7 & -28 & 14 \end{bmatrix}$,

$AC=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $CA=\begin{bmatrix} 0 & 0 & 0 \\ 2 & -8 & 4 \\ 4 & -16 & 8 \end{bmatrix}$, $A(2B-3C)=\begin{bmatrix} 30 & -28 \\ -30 & 28 \end{bmatrix}$

2. (a) $\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}$, a 与 b 任意　(b) $\begin{bmatrix} -2a & a \\ -2b & b \end{bmatrix}$, a 与 b 任意

3. (a) $a=9$, $b=6$, $c=1$, $d=5$　(b) $a=1$, $b=6$, $c=0$, $d=-2$

4. (a) $\begin{bmatrix} -9 & -2 & -10 \\ 6 & 14 & 8 \\ -7 & 5 & -5 \end{bmatrix}$ (b) $\begin{bmatrix} -3 & 5 & -4 \\ 0 & 3 & 24 \\ 12 & -27 & 0 \end{bmatrix}$

6. $A^n=\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

7. $A^n=\begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$

8. $A^n=\begin{bmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 \\ -100 & 1 \end{bmatrix}$

10. $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, 此处 b 与 c 任意, a 满足 $a^2 = -bc$

11. $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, 此处 a 任意

12. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ 与 $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, 此处 b 与 c 任意, a 为满足条件 $a^2 = 1 - bc$ 的任意数

13. $\boldsymbol{C} = \begin{bmatrix} \frac{15}{2} & \frac{13}{2} \\ 8 & 7 \end{bmatrix}$, $\boldsymbol{D} = \begin{bmatrix} \frac{33}{4} & \frac{19}{4} \\ \frac{43}{4} & \frac{25}{4} \end{bmatrix}$

14. (b) $(\boldsymbol{A}+\boldsymbol{B})^2 = \boldsymbol{A}^2 + \boldsymbol{AB} + \boldsymbol{BA} + \boldsymbol{B}^2$; $(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B}) = \boldsymbol{A}^2 + \boldsymbol{BA} - \boldsymbol{AB} + \boldsymbol{B}^2$

    (c) 对那些可交换的矩阵

4.20 习题

1. $(x, y, z) = (\frac{8}{5}, -\frac{7}{5}, \frac{8}{5})$
2. 无解
3. $(x, y, z) = (1, -1, 0) + t(-3, 4, 1)$
4. $(x, y, z) = (1, -1, 0) + t(-3, 4, 1)$
5. $(x, y, z, u) = (1, 1, 0, 0) + t(1, 14, 5, 0)$
6. $(x, y, z, u) = (1, 8, 0, -4) + t(2, 7, 3, 0)$
7. $(x, y, z, u, v) = t_1(-1, 1, 0, 0, 0) + t_2(-1, 0, 3, -3, 1)$
8. $(x, y, z, u) = (1, 1, 1, -1) + t_1(-1, 3, 7, 0) + t_2(4, 9, 0, 7)$
9. $(x, y, z) = (\frac{4}{3}, \frac{2}{3}, 0) + t(5, 1, -3)$
10. (a) $(x, y, z, u) = (1, 6, 3, 0) + t_1(4, 11, 7, 0) + t_2(0, 0, 0, 1)$

    (b) $(x, y, z, u) = (\frac{3}{11}, 4, \frac{19}{11}, 0) + t(4, -11, 7, 22)$

12. $\begin{bmatrix} -1 & 2 & 1 \\ 5 & -8 & -6 \\ -3 & 5 & 4 \end{bmatrix}$

13. $\begin{bmatrix} \frac{-5}{3} & \frac{2}{3} & \frac{4}{3} \\ -1 & 0 & 1 \\ \frac{7}{3} & \frac{-1}{3} & \frac{-5}{3} \end{bmatrix}$

14. $\begin{bmatrix} 14 & 8 & 3 \\ 8 & 5 & 2 \\ 3 & 2 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

16. $\begin{bmatrix} 0 & \frac{1}{2} & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ -3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 9 & 0 & -3 & 0 & 1 & 0 \end{bmatrix}$

4.21 关于矩阵的综合性习题

3. $\boldsymbol{P} = \begin{bmatrix} 2 & 1 \\ 5 & -1 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 与 $\begin{bmatrix} a & b \\ c & 1-a \end{bmatrix}$, 此处 b 与 c 任意, a 为二次方程 $a^2 - a + bc = 0$ 的任意解

10. (a) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$,

$\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$, $\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$

5.8 习题

1. (a) 6 (b) 76 (c) $a^3 - 4a$
2. (a) 1 (b) 1 (c) 1
3. (b) $(b-a)(c-a)(c-b)(a+b+c)$ 与 $(b-a)(c-a)(c-b)(ab+ac+bc)$
5. (a) -8 (b) $(b-a)(c-a)(d-a)(c-b)(d-b)(d-c)$
   (c) $(b-a)(c-a)(d-a)(c-b)(d-b)(d-c)(a+b+c+d)$
   (d) $a(a^2-4)(a^2-16)$ (e) -160
7. $\det \boldsymbol{A} = 16$, $\det(\boldsymbol{A}^{-1}) = \frac{1}{16}$, $\boldsymbol{A}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{3}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & \frac{1}{2} & -\frac{3}{4} & \frac{1}{8} \\ 0 & 0 & \frac{1}{2} & -\frac{3}{4} \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}$

10. $F' = \begin{vmatrix} f_1' & f_2' & f_3' \\ g_1 & g_2 & g_3 \\ h_1 & h_2 & h_3 \end{vmatrix} + \begin{vmatrix} f_1 & f_2 & f_3 \\ g_1' & g_2' & g_3' \\ h_1 & h_2 & h_3 \end{vmatrix} + \begin{vmatrix} f_1 & f_2 & f_3 \\ g_1 & g_2 & g_3 \\ h_1' & h_2' & h_3' \end{vmatrix}$

11. (b) 若 $F = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{vmatrix}$ 则 $F' = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1''' & f_2''' & f_3''' \end{vmatrix}$

5.15 习题

6. $\det \boldsymbol{A} = (\det \boldsymbol{B})(\det \boldsymbol{D})$
7. (a) 线性无关 (b) 线性无关 (c) 线性相关

5.20 习题

1. (a) $\begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 & 1 \\ -6 & 3 & 5 \\ -4 & -2 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 109 & 113 & -41 & -13 \\ -40 & -92 & 74 & 16 \\ -41 & -79 & 7 & 47 \\ -50 & 38 & 16 & 20 \end{bmatrix}$

2. (a) $-\frac{1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ (b) $\frac{1}{8}\begin{bmatrix} 2 & -6 & -4 \\ -1 & 3 & -2 \\ 1 & 5 & 2 \end{bmatrix}$

   (c) $\frac{1}{306}\begin{bmatrix} 109 & -40 & -41 & -50 \\ 113 & -92 & -79 & 38 \\ -41 & 74 & 7 & 16 \\ -13 & 16 & 47 & 20 \end{bmatrix}$

3. (a) $\lambda = 1$ (b) $\lambda = 0, \lambda = \pm 3$ (c) $\lambda = 3, \lambda = \pm \mathrm{i}$

5. (a) $x = 0, y = 1, z = 2$ (b) $x = 1, y = 1, z = -1$

5.22 关于行列式的综合性习题

1. (b) $\det\begin{bmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{bmatrix} = 0;\det\begin{bmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{bmatrix} = 0$

   (c) $\det\begin{bmatrix} (x - x_1)^2 + (y - y_1)^2 & (x - x_1) & (y - y_1) \\ (x_2 - x_1)^2 + (y_2 - y_1)^2 & (x_2 - x_1) & (y_2 - y_1) \\ (x_3 - x_1)^2 + (y_3 - y_1)^2 & (x_3 - x_1) & (y_3 - y_1) \end{bmatrix} = 0;$

   $\det\begin{bmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{bmatrix} = 0.$

2. (a) $\operatorname{cof} \boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ -x & 1 & 0 \\ x^2 - y & -x & 1 \end{bmatrix}$

5. $6(x - x^2 + \mathrm{e}^{3x} - \mathrm{e}^{2x})$

6. $a = -7, b = c = 12$

6.4 习题

1. (b) $a\lambda_1 + b\lambda_2$
7. 非零常数多项式
8. 特征函数：$f(t) = Ct^{\lambda}$，此处 $C \neq 0$
9. 特征函数：$f(t) = C\mathrm{e}^{t/\lambda}$，此处 $C \neq 0$
10. 特征函数：$f(t) = C\mathrm{e}^{t^2/(2\lambda)}$，此处 $C \neq 0$

11. 属于 $\lambda = 0$ 的特征向量是极限为 $a \neq 0$ 的所有常数序列. 属于 $\lambda = -1$ 的特征向量是极限为 $a = 0$ 的所有非常数序列.

6.10 习题

特征值	特征向量	$\dim E(\lambda)$
1. (a) 1, 1	$(a, b) \neq (0, 0)$	2
(b) 1, 1	$t(1, 0)$, $t \neq 0$	1
(c) 1, 1	$t(0, 1)$, $t \neq 0$	1
(d) 2	$t(1, 1)$, $t \neq 0$	1
0	$t(1, -1)$, $t \neq 0$	1
2. $1 + \sqrt{ab}$	$t(\sqrt{a}, \sqrt{b})$, $t \neq 0$	1
$1 - \sqrt{ab}$	$t(\sqrt{a}, -\sqrt{b})$, $t \neq 0$	1

3. 若纯量域是全体实数的集合 $\mathbb{R}$, 则只有当 $\sin\theta = 0$ 时实特征值存在, 此时有两个相同的特征值 $\lambda_1 = \lambda_2 = \cos\theta$, 其中 $\cos\theta = 1$ 或 -1. 在这一情形下每一个非零向量都是特征向量, 因此 $\dim E(\lambda_1) = \dim E(\lambda_2) = 2$. 若纯量域是全体复数的集合 $\mathbb{C}$, 则特征值为 $\lambda_1 = \cos\theta + i\sin\theta$ 和它的共轭复数 $\lambda_2 = \cos\theta - i\sin\theta$. 若 $\sin\theta = 0$, 则 λ_1 与 λ_2 是两个相同的实数. 若 $\sin\theta \neq 0$, 则 λ_1 与 λ_2 是一对不相等的共轭复数; 属于 λ_1 的特征向量为 $t(i, 1)$, 此处 $t \neq 0$; 属于 λ_2 的特征向量为 $t(1, i)$, 此处 $t \neq 0$; $\dim E(\lambda_1) = \dim E(\lambda_2) = 1$.

4. $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, 此处 b 与 c 任意, a 是满足条件 $a^2 = 1 - bc$ 的任意数

5. 设 $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, 且令 $\Delta = (a-d)^2 + 4bc$. 若 $\Delta > 0$, 则特征值为两个不同的实数. 若 $\Delta = 0$, 则特征值为两个相同的实数, 若 $\Delta < 0$, 则特征值是一对共轭复数.

6. $a = b = c = d = e = f = 1$

特征值	特征向量	$\dim E(\lambda)$
7. (a) 1, 1, 1	$t(0, 0, 1)$, $t \neq 0$	1
(b) 1	$t(1, -1, 0)$, $t \neq 0$	1
2	$t(3, 3, -1)$, $t \neq 0$	1
21	$t(1, 1, 6)$, $t \neq 0$	1
(c) 1	$t(3, -1, 3)$, $t \neq 0$	1
2, 2	$t(2, 2, -1)$, $t \neq 0$	1

8. 对每一个矩阵, 特征值都是 1, 1, -1, -1.

6.12 习题

2. (a) 特征值为 1, 3; $\boldsymbol{C} = \begin{bmatrix} -2c & 0 \\ c & d \end{bmatrix}$, 此处 $cd \neq 0$

(b) 特征值为 6, -1; $\boldsymbol{C} = \begin{bmatrix} 2a & b \\ 5a & -b \end{bmatrix}$, 此处 $ab \neq 0$

(c) 特征值为 3, 3; 若存在非奇异矩阵 $\boldsymbol{C}$, 则 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = 3\boldsymbol{I}$, $\boldsymbol{A}\boldsymbol{C} = 3\boldsymbol{C}$, $\boldsymbol{A} = 3\boldsymbol{I}$

(d) 特征值为 1, 1; 若存在非奇异矩阵 $\boldsymbol{C}$, 则 $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{I}$, $\boldsymbol{A}\boldsymbol{C} = \boldsymbol{C}$, $\boldsymbol{A} = \boldsymbol{I}$

3. $\boldsymbol{C} = \boldsymbol{A}^{-1}\boldsymbol{B}$

4. (a) 特征值为 1, 1, −1; 特征向量为 (1,0,1), (0,1,0), (1,0,−1);

$$C=\begin{bmatrix}1&0&1\\0&1&0\\1&0&-1\end{bmatrix}$$

(b) 特征值为 2, 2, 1; 特征向量为 (1,0,−1), (0,1,−1), (1,−1,1);

$$C=\begin{bmatrix}1&0&1\\0&1&-1\\-1&-1&1\end{bmatrix}$$

5. (a) 特征值为 2, 2; 特征向量为 $t(1,0)$, 此处 $t\neq 0$. 若 $C=\begin{bmatrix}a&b\\-b&0\end{bmatrix}$, 此处 $b\neq 0$, 则

$$C^{-1}AC=\begin{bmatrix}2&0\\1&2\end{bmatrix}$$

(b) 特征值为 3, 3; 特征向量为 $t(1,1)$, 此处 $t\neq 0$. 若 $C=\begin{bmatrix}a&b\\a+b&b\end{bmatrix}$, 此处 $b\neq 0$, 则

$$C^{-1}AC=\begin{bmatrix}3&0\\1&3\end{bmatrix}$$

6. 特征值为 1, 1, 1; 特征向量为 $t(1,-1,-1)$, 此处 $t\neq 0$

6.14 习题

1. $A^{-1}=2I-A, A^n=nA-(n-1)I$
2. $A^{-1}=\frac{3}{2}I-\frac{1}{2}A, A^n=(2^n-1)A-(2^n-2)I$
3. $A^{-1}=A, A^n=\frac{1+(-1)^n}{2}I+\frac{1-(-1)^n}{2}A$
4. $A^{-1}=A, A^n=\frac{1+(-1)^n}{2}I+\frac{1-(-1)^n}{2}A$
5. $A^n=O$, 若 $n\geqslant 3$
6. $A^n=A$, 若 $n\geqslant 1$
7. $A^3=4A^2-5A+2I$; $A^n=\begin{bmatrix}2^n&0&0\\0&1&0\\0&n&1\end{bmatrix}$

6.16 关于特征值与特征向量的综合性习题

1. (a) 特征值 $\lambda_1=-1$, 特征向量为 $t(1,0,0)$, 此处 $t\neq 0$.
   特征值 $\lambda_2=4$, 特征向量为 $t(-2,1,0)$, 此处 $t\neq 0$.
   特征值 $\lambda_3=0$, 特征向量为 $t(1,-2,1)$, 此处 $t\neq 0$.

   (b) $C=\begin{bmatrix}1&-2&1\\0&1&-2\\0&0&1\end{bmatrix}$

2. $C=\begin{bmatrix}1&a&0\\-1&-a&a\\-1&a&-a\end{bmatrix}$, 此处 a 为任意非零数

3. (a) 特征值为 $\lambda_1=1, \lambda_2=\lambda_3=3$. 属于 λ_1 的特征向量为 $a(1,-1,-1)$, 此处 $a\neq 0$, 属于 λ_2 和 λ_3

的特征向量为 $a(-1,2,0)$, 此处 $a\neq 0$.

(b) $C=\begin{bmatrix}1 & -1 & -1\\ -1 & 2 & 0\\ -1 & 0 & 3\end{bmatrix}$

4. (b) $C=\begin{bmatrix}1 & 0 & x\\ 1 & 0 & 0\\ 1 & 2x & y\end{bmatrix}$, 此处 x 与 y 为任意数且 $x\neq 0$.

(c) $D=\begin{bmatrix}0 & x & c\\ 0 & 0 & c\\ 2x & y & c\end{bmatrix}$, 此处 x 与 y 为任意数且 $xc\neq 0$.

5. 特征值为 $2+i$, $2-i$; 正交特征向量为 $(1,i)$, $(1,-i)$

6. $\begin{bmatrix}\frac{7}{2} & -\frac{1}{2} & -\frac{1}{2}\\ 1 & 2 & -1\\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{bmatrix}$

7. (a) $J^2=nJ$, $J^3=n^2J$, $\cdots$, $J^n=n^{n-1}J$ (b) $\det(\lambda I-J)=\lambda^{n-1}(\lambda-n)$

(c) 正交特征向量组为 $(1,1,1,1)$, $(1,-1,1,-1)$, $(-1,1,1,-1)$, $(-1,-1,1,1)$. 第一个是属于 $\lambda=4$ 的特征向量, 其余三个是属于 $\lambda=0$ 的特征向量.

(d) 除了一个特征值为 nt 外, 其余特征值都等于零.

8. 除了一个特征值为 $nt+a-t$ 外, 其余特征值都等于 $a-t$.

9. $A^2=n^2I$; $A^3=n^2I$; $A^{-1}=\frac{1}{n^2}A$

10. (a) $a=\frac{1}{2}(1+(-1)^n)$, $b=\frac{1}{2}(1-(-1)^n)$ (b) 特征值为 1, 1, -1, -1

(c) $C=\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & 0 & 1 & 0\\ 1 & 0 & -1 & 0\\ 1 & 1 & -1 & -1\end{bmatrix}$

11. (a) $A^4=16I$

(b) $\det(\lambda I-A)=\lambda^4-16$; 特征值为 2, 2, -2, -2

12. (a) $A^2=2A+3I$

(b) $A^{-1}=\frac{1}{3}(A-2I)=\frac{1}{3}\begin{bmatrix}-2 & 1 & 1 & 1\\ 1 & -2 & 1 & 1\\ 1 & 1 & -2 & 1\\ 1 & 1 & 1 & -2\end{bmatrix}$

(c) 特征值为 -1, -1, -1, 3

(d) 特征向量为: $(1,0,0,-1)$, $(0,1,0,-1)$, $(0,0,1,-1)$, $(1,1,1,1)$

7.4 习题

3. (b) 若 n 为偶数, 则 T^n 是 Hermite 矩阵, 若 n 为奇数, 则 T^n 是斜 Hermite 矩阵.

6. (d) $Q(x+ty)=Q(x)+t\bar{t}Q(y)+\bar{t}(T(x),y)+t(T(y),x)$

7.10 习题

1. (a) 对称矩阵与 Hermite 矩阵

(b) 不属于这四种类型

(c) 斜对称矩阵

(d) 斜对称矩阵与斜 Hermite 矩阵

4. (b) $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$

5. 特征值为 $\lambda_1 = 0, \lambda_2 = 25$; 标准正交特征向量组为 $\boldsymbol{u}_1 = \frac{1}{5}(4,-3)$, $\boldsymbol{u}_2 = \frac{1}{5}(3,4)$;

$\boldsymbol{C} = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix}$

6. 特征值为 $\lambda_1 = 2i, \lambda_2 = -2i$; 标准正交特征向量组为 $\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1,-i)$,

$\boldsymbol{u}_2 = \frac{1}{\sqrt{2}}(1,i)$; $\boldsymbol{C} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$

7. 特征值为 $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = -4$; 标准正交特征向量组为 $\boldsymbol{u}_1 = \frac{1}{\sqrt{10}}(1,0,3)$, $u_2 = \frac{1}{\sqrt{14}}(3,2,-1)$,

$\boldsymbol{u}_3 = \frac{1}{\sqrt{35}}(3,-5,-1)$; $\boldsymbol{C} = \begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{14}} & \frac{3}{\sqrt{35}} \\ 0 & \frac{2}{\sqrt{14}} & \frac{-5}{\sqrt{35}} \\ \frac{3}{\sqrt{10}} & \frac{-1}{\sqrt{14}} & \frac{-1}{\sqrt{35}} \end{bmatrix}$

8. 特征值为 $\lambda_1 = 1, \lambda_2 = 6, \lambda_3 = -4$; 标准正交特征向量组为 $\boldsymbol{u}_1 = \frac{1}{5}(0,4,-3)$,

$\boldsymbol{u}_2 = \frac{1}{\sqrt{50}}(5,3,4)$, $\boldsymbol{u}_3 = \frac{1}{\sqrt{50}}(5,-3,-4)$; $\boldsymbol{C} = \frac{1}{\sqrt{50}}\begin{bmatrix} 0 & 5 & 5 \\ 4\sqrt{2} & 3 & -3 \\ -3\sqrt{2} & 4 & -4 \end{bmatrix}$

9. (a), (b) 与 (c) 为酉矩阵; (b) 与 (c) 为正交矩阵

11. (a) 特征值为 $\lambda_1 = ia, \lambda_2 = -ia$; 标准正交特征向量组为 $\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1,i)$,

$\boldsymbol{u}_2 = \frac{1}{\sqrt{2}}(1,-i)$; $\boldsymbol{C} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$

7.14 习题

1. (a) $\boldsymbol{A} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ (b) $\lambda_1 = 0, \lambda_2 = 5$

(c) $\boldsymbol{u}_1 = \frac{1}{\sqrt{5}}(1,-2), \boldsymbol{u}_2 = \frac{1}{\sqrt{5}}(2,1)$; (d) $\boldsymbol{C} = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$

2. (a) $\boldsymbol{A} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$ (b) $\lambda_1 = \frac{1}{2}, \lambda_2 = -\frac{1}{2}$

(c) $\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}(1,1), \boldsymbol{u}_2 = \frac{1}{\sqrt{2}}(1,-1)$; (d) $\boldsymbol{C} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

3. (a) $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ (b) $\lambda_1 = \sqrt{2}, \lambda_2 = -\sqrt{2}$

(c) $u_1 = t(1+\sqrt{2}, 1)$, $u_2 = t(-1, 1+\sqrt{2})$, 此处 $t = 1/\sqrt{4+2\sqrt{2}}$.

(d) $\boldsymbol{C} = t\begin{bmatrix} 1+\sqrt{2} & -1 \\ 1 & 1+\sqrt{2} \end{bmatrix}$, 此处 $t = 1/\sqrt{4+2\sqrt{2}}$

4. (a) $\boldsymbol{A}=\begin{bmatrix} 34 & -12 \\ -12 & 41 \end{bmatrix}$ (b) $\lambda_1=50, \lambda_2=25$

(c) $\boldsymbol{u}_1=\frac{1}{5}(3,-4), \boldsymbol{u}=\frac{1}{5}(4,3)$; (d) $\boldsymbol{C}=\frac{1}{5}\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$

5. (a) $\boldsymbol{A}=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ (b) $\lambda_1=0, \lambda_2=\frac{3}{2}, \lambda_3=-\frac{1}{2}$

(c) $\boldsymbol{u}_1=\frac{1}{\sqrt{3}}(1,-1,-1), \boldsymbol{u}_2=\frac{1}{\sqrt{6}}(2,1,1), \boldsymbol{u}_3=\frac{1}{\sqrt{2}}(0,1,-1)$

(d) $\boldsymbol{C}=\frac{1}{\sqrt{6}}\begin{bmatrix} \sqrt{2} & 2 & 0 \\ -\sqrt{2} & 1 & \sqrt{3} \\ -\sqrt{2} & 1 & -\sqrt{3} \end{bmatrix}$

6. (a) $\boldsymbol{A}=\begin{bmatrix} 2 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$ (b) $\lambda_1=1, \lambda_2=3, \lambda_3=-2$

(c) $\boldsymbol{u}_1=(0,1,0), \boldsymbol{u}_2=\frac{1}{\sqrt{5}}(2,0,1), \boldsymbol{u}_3=\frac{1}{\sqrt{5}}(1,0,-2)$

(d) $\boldsymbol{C}=\frac{1}{\sqrt{5}}\begin{bmatrix} 0 & 2 & 1 \\ \sqrt{5} & 0 & 0 \\ 0 & 1 & -2 \end{bmatrix}$

7. (a) $\boldsymbol{A}=\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ (b) $\lambda_1=\lambda_2=-1, \lambda_3=8$

(c) $\boldsymbol{u}_1=\frac{1}{\sqrt{2}}(1,0,-1), \boldsymbol{u}_2=\frac{1}{3\sqrt{2}}(-1,4,-1), \boldsymbol{u}_3=\frac{1}{3}(2,1,2)$

(d) $\boldsymbol{C}=\frac{1}{3\sqrt{2}}\begin{bmatrix} 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \\ -3 & -1 & 2\sqrt{2} \end{bmatrix}$

8. 椭圆; 中心坐标为 $(0,0)$

9. 双曲线; 中心坐标为 $(-\frac{5}{2},-\frac{5}{2})$

10. 抛物线; 顶点坐标为 $(\frac{5}{16},-\frac{15}{16})$

11. 椭圆; 中心坐标为 $(0,0)$

12. 椭圆; 中心坐标为 $(6,-4)$

13. 抛物线; 顶点坐标为 $(\frac{2}{25},\frac{11}{25})$

14. 椭圆; 中心坐标为 $(0,0)$

15. 抛物线; 顶点坐标为 $(\frac{3}{4},\frac{3}{4})$

16. 椭圆; 中心坐标为 $(-1,\frac{1}{2})$

17. 双曲线; 中心坐标为 $(0,0)$

18. 两条相交直线

19. -14

7.20 习题

8. $a=\pm\frac{1}{3}\sqrt{3}$

13. (a), (b) 与 (e)

7.22 习题

2. (a) 对称变换 (b) 都不是 (c) 对称变换 (d) 对称变换

8.3 习题

1. $y = e^{3x} - x^{2x}$
2. $y = \frac{2}{3}x^2 + \frac{1}{3}x^5$
3. $y = 4\cos x - 2\cos^2 x$
4. 四倍于开始时的总量
5. $f(x) = Cx^n$ 或 $f(x) = Cx^{1/n}$
6. (b) $y = e^{4x} - e^{-x^3/3}$
7. $y = -1 \pm \sqrt{2 + ce^{-x^2}}$, 此处 $c \geqslant -2$
8. $y = c_1e^{2x} + c_2e^{-2x}$
9. $y = c_1\cos 2x + c_2\sin 2x$
10. $y = e^x(c_1\cos 2x + c_2\sin 2x)$
11. $y = e^{-x}(c_1 + c_2x)$
12. (a) $y = 10e^{3x} - 5e^{-x}$ (b) $A = -\frac{1}{2}, B = 2, C = -1$
13. $k = n^2\pi^2$; $f_k(x) = C\sin n\pi x (n = 1, 2, 3, \cdots)$
15. (a) $y'' - y = 0$
    (b) $y'' - 4y' + 4y = 0$
    (c) $y'' + y' + \frac{5}{4}y = 0$
    (d) $y'' + 4y = 0$
    (e) $y'' - y = 0$
16. $y = \frac{1}{3}\sqrt{6}, y'' = -12y = -4\sqrt{6}$

8.9 习题

1. $y = c_1 + c_2e^{-x} + c_3e^{3x}$
2. $y = c_1 + c_2e^{x} + c_3e^{-x}$
3. $y = c_1 + (c_2 + c_3x)e^{-2x}$
4. $y = (c_1 + c_2x + c_3x^2)e^{x}$
5. $y = (c_1 + c_2x + c_3x^2 + c_4x^3)e^{-x}$
6. $y = c_1e^{2x} + c_2e^{-2x} + c_3\cos 2x + c_4\sin 2x$
7. $y = e^{\sqrt{2x}}(c_1\cos\sqrt{2x} + c_2\sin\sqrt{2}x) + e^{-\sqrt{2}x}(c_3\cos\sqrt{2}x + c_4\sin\sqrt{2}x)$
8. $y = c_1e^{x} + e^{-x/2}(c_2\cos\frac{1}{2}\sqrt{3}x + c_3\sin\frac{1}{2}\sqrt{3}x)$
9. $y = e^{-x}\{(c_1 + c_2x)\cos x + (c_3 + c_4x)\sin x\}$
10. $y = (c_1 + c_2x)\cos x + (c_3 + c_4x)\sin x$
11. $y = c_1 + c_2x + (c_3 + c_4x)\cos\sqrt{2}x + (c_5 + c_6x)\sin\sqrt{2}x$
12. $y = c_1 + c_2x + (c_3 + c_4x)\cos 2x + (c_5 + c_6x)\sin 2x$
13. $f(x) = \frac{1}{2m^2}(e^{mx} - \cos mx - \sin mx)$
15. (a) $y^{(4)} - 5y'' + 4y = 0$
    (b) $y''' + 6y'' + 12y' + 8y = 0$
    (c) $y^{(4)} - 2y''' + y'' = 0$
    (d) $y^{(4)} - 2y''' + y'' = 0$

(e) $y^{(5)} - 2y^{(4)} + y''' = 0$
(f) $y^{(4)} + 8y''' + 33y'' + 68y' + 52y = 0$
(g) $y^{(4)} - 2y'' + y = 0$
(h) $y^{(6)} + 4y'' = 0$

8.15 习题

1. $y_1 = -2x - x^2 - \frac{1}{3}x^3$
2. $y_1 = \frac{1}{4}xe^{2x}$
3. $y_1 = (x - \frac{4}{3})e^x$
4. $y_1 = \frac{1}{3}\sin x$
5. $y_1 = \frac{1}{2}x^2e^x + e^{2\lambda}$
6. $y_1 = \frac{1}{2}xe^x$
7. $y_1 = x\cosh x$
8. $y_1 = \frac{1}{24}x^4e^{-x}$
9. $50y_1 = (11 - 5x)e^x\sin 2x + (2 - 10x)e^x\cos 2x$
10. $y_1 = -(\frac{5}{8}x + \frac{3}{8}x^2 + \frac{1}{12}x^3)e^{-x}$
12. (c) $y_1 = \frac{x^m e^{\alpha x}}{p_A^{(m)}(\alpha)}$
15. (b) $2D$ (c) $3D^2$ (d) nD^{n-1}
16. $u(x) = 6(e^{4x} - e^{-x})/5; v(x) = e^x - e^{-5\lambda}$
17. $u(x) = \frac{1}{2}e^{2x-\pi}\sin 5x; v(x) = \frac{5}{6}e^{-2x-\pi}\sin 3x$
18. $u(x) = e^{-x^2}; Q(x) = 4x^2 + 2$
20. $y = (A + Bx^3)e^{\lambda} + (x^2 - 2x + 2)e^{2x}$
21. $y = Ax^{1/2} + Bx^{-1/2}$
22. $y = A(x^2 - 2) + B/x$
23. $y = x^{-2}[A + B(x-1)^3 + \frac{1}{9}x^3 + \frac{2}{3}x^2 - \frac{7}{6}x + \frac{1}{2} - (x-1)^3\ln|x-1|]$
24. (a) $P(x) = -5/x$; $Q(x) = 9/x^2$
    (b) $y = c_1x^3 + c_2x^3\ln x + x^4$
25. 每一个 $\lambda > 0$ 都是特征值; 相应的特征函数为 $f(t) = Ce^{t/\sqrt{\lambda}}$, 此处 $C \neq 0$
27. 仅有的可能特征值为 $\lambda = 0$, $\lambda = -n^2$, 此处 $n = 1, 2, 3, \cdots$. 属于 $\lambda = 0$ 的特征函数为 $f(t) = a$, 此处 $a \neq 0$. 属于 $\lambda = -n^2$ 的特征函数为 $f(t) = a\cos nt$, 此处 $a \neq 0$.
28. 属于 $\lambda = 0$ 的特征函数为 $f(t) = at$, 此处 $a \neq 0$. 属于 $\lambda > 0$ 的特征函数为 $f(t) = c\sinh\sqrt{\lambda}t$, 此处 $c \neq 0$. 若 $\lambda = 1$, 我们也有 $f(t) = c_1e^t + c_2e^{-t}$, 此处 c_1 与 c_2 不同时为零. 属于 $\lambda < 0$ 的特征函数为 $f(t) = a\sin\sqrt{-\lambda}t$, 此处 $a \neq 0$.
29. 属于 $\lambda = 0$ 的特征函数为 $f(t) = a(2+t)$, 此处 $a \neq 0$. 属于 $\lambda > 0$ 的特征函数为 $f(t) = a(e^{\sqrt{\lambda}t/2} + \frac{\sqrt{\lambda}-1}{\sqrt{\lambda}+1}e^{-\sqrt{\lambda}t/2})$, 此处 $a \neq 0$. 属于 $\lambda < 0$ 的特征函数为 $f(t) = a(\cos\sqrt{-\lambda}t + \frac{1}{\sqrt{-\lambda}}\sin\sqrt{-\lambda}t)$, 此处 $a \neq 0$.
30. 属于 $\lambda = 0$ 的特征函数为 $f(t) = c$, 此处 $c \neq 0$. 属于 $\lambda > 0$ 的特征函数为 $f(t) = c(e^{-\sqrt{\lambda}}e^{\sqrt{\lambda}t} +$

$e^{-\sqrt{\lambda}t})$, 此处 $c \neq 0$. 属于 $\lambda < 0$ 的特征函数分为两类: 若 $\lambda = -4n^2\pi^2$, 此处 n 为非零整数, 则属于 λ 的特征函数为 $f(t) = c_1 \cos 2\pi nt + c_2 \sin 2\pi nt$, 此处 c_1 与 c_2 不同时为零; 若 $\lambda < 0$ 且 $\lambda \neq -4n^2\pi^2$, 则

$$f(t) = c\left(\frac{\sin\sqrt{-\lambda}}{1-\cos\sqrt{-\lambda}}\cos\sqrt{-\lambda}t + \sin\sqrt{-\lambda}t\right),$$

此处 $c \neq 0$.

9.4 习题

3. (b) $(P_k)' = \sum_{m=0}^{k-1} P^m P' P^{k-1-m}$

9.11 习题

1. (a) $\boldsymbol{A}^{-1} = 2\boldsymbol{I} - \boldsymbol{A}, \boldsymbol{A}^n = n\boldsymbol{A} - (n-1)\boldsymbol{I}$

   (b) $e^{t\boldsymbol{A}} = e^t(1-t)\boldsymbol{I} + te^t\boldsymbol{A} = e^t\begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}$

2. (a) $\boldsymbol{A}^{-1} = \frac{3}{2}\boldsymbol{I} - \frac{1}{2}\boldsymbol{A}, \boldsymbol{A}^n = (2^n-1)\boldsymbol{A} - (2^n-2)\boldsymbol{I}$

   (b) $e^{t\boldsymbol{A}} = (2e^t - e^{2t})\boldsymbol{I} + (e^{2t} - e^t)\boldsymbol{A} = \begin{bmatrix} e^t & 0 \\ e^{2t} - e^t & e^{2t} \end{bmatrix}$

3. (a) $\boldsymbol{A}^{-1} = A, \boldsymbol{A}^n = \frac{1+(-1)^n}{2}\boldsymbol{I} + \frac{1-(-1)^n}{2}\boldsymbol{A}$

   (b) $e^{t\boldsymbol{A}} = (\cosh t)\boldsymbol{I} + (\sinh t)\boldsymbol{A} = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}$

4. (a) $\boldsymbol{A}^{-1} = \boldsymbol{A}, \boldsymbol{A}^n = \frac{1+(-1)^n}{2}\boldsymbol{I} + \frac{1-(-1)^n}{2}\boldsymbol{A}$

   (b) $e^{t\boldsymbol{A}} = (\cosh t)\boldsymbol{I} + (\sinh t)\boldsymbol{A} = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^t \end{bmatrix}$

5. (a) $\boldsymbol{A}^n = \boldsymbol{O}$, 若 $n \geqslant 3$

   (b) $e^{t\boldsymbol{A}} = \boldsymbol{I} + t\boldsymbol{A} + \frac{1}{2}t^2\boldsymbol{A}^2 = \begin{bmatrix} 1 & t & t+\frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}$

6. (a) $\boldsymbol{A}^n = \boldsymbol{A}$, 若 $n \geqslant 1$

   (b) $e^{t\boldsymbol{A}} = \boldsymbol{I} + (e^t - 1)\boldsymbol{A} = \begin{bmatrix} 1 & e^t-1 & e^t-1 \\ 0 & e^t & e^t-1 \\ 0 & 0 & 1 \end{bmatrix}$

7. (a) $\boldsymbol{A}^3 = 4\boldsymbol{A}^2 - 5\boldsymbol{A} + 2\boldsymbol{I}; \boldsymbol{A}^n = \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & n & 1 \end{bmatrix}$

   (b) $e^{t\boldsymbol{A}} = \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & te^t & e^t \end{bmatrix}$

8. (b) $e^{tA} = e^{ut}\begin{bmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{bmatrix}$

10. $e^{A(t)} = I + (e-1)A(t); (e^{A(t)})^t = (e-1)A^t(t) = \begin{bmatrix} 0 & e-1 \\ 0 & 0 \end{bmatrix}$;

$e^{A(t)}A^t(t) = \begin{bmatrix} 0 & e \\ 0 & 0 \end{bmatrix}$; $A^t(t)e^{A(t)} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

11. $e^{tA} = I + tA + \frac{1}{2}t^2A^2$

13. $e^Ae^B = \begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix}$; $e^Be^A = \begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$; $e^{A+B} = \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$

15. (b) $e^{tA} = (\cosh nt)I + \frac{1}{n}(\sinh nt)A$

16. (b) $f(t) = \frac{1}{5}e^{4t} + \frac{1}{5}e^{-t}; g(t) = \frac{1}{5}e^{4t} - \frac{1}{5}e^{-t}$.

9.14 习题

1. $e^{tA} = \frac{1}{2}(3e^t - e^{3t})I + \frac{1}{2}(e^{3t} - e^t)A$
2. $e^{tA} = (\cosh\sqrt{5}t)I + \frac{1}{\sqrt{5}}(\sinh\sqrt{5}t)A$
3. $e^{tA} = \frac{1}{2}e^t\{t^2 - 2t + 2)I + (-2t^2 + 2t)A + t^2A^2\}$
4. $e^{tA} = (3e^{-t} - 3e^{-2t} + e^{-3t})I + (\frac{5}{2}e^{-t} - 4e^{-2t} + \frac{3}{2}e^{-3t})A + (\frac{1}{2}e^{-t} - e^{-2t} + \frac{1}{2}e^{-3t})A^2$
5. $e^{tA} = (4e^t - 3e^{2t} + 2te^{2t})I + (4e^{2t} - 3te^{2t} - 4e^t)A + (e^t - e^{2t} + te^{2t})A^2$
6. $e^{tA} = (4e^t - 6e^{2t} + 4e^{3t} - e^{4t})I + (-\frac{13}{3}e^t + \frac{19}{2}e^{2t} - 7e^{3t} + \frac{11}{6}e^{4t})A$

   $+(\frac{3}{2}e^t - 4e^{2t} + \frac{7}{2}e^{3t} - e^{4t})A^2 + (-\frac{1}{6}e^t + \frac{1}{2}e^{2t} - \frac{1}{2}e^{3t} + \frac{1}{6}e^{4t})A^3$
7. (b) $e^{tA} = \frac{1}{6}e^{\lambda t}\{6 - 6\lambda t + 3\lambda^2t^2 - \lambda^3t^3)I + (6t - 6\lambda t^2 + 3\lambda^2t^3)A$

   $+(3t^2 - 3\lambda t^3)A^2 + t^3A^3\}$
9. $y_1 = c_1\cosh\sqrt{5}t + \frac{c_1+2c_2}{\sqrt{5}}\sinh\sqrt{5}t, y_2 = c_2\cosh\sqrt{5}t + \frac{2c_1-c_2}{\sqrt{5}}\sinh\sqrt{5}t$
10. $y_1 = e^t(\cos 3t - \sin 3t), y_2 = e^t(\cos 3t - 3\sin 3t)$
11. $y_1 = e^{2t} + 4te^{2t}, y_2 = -2e^t + e^{2t} + 4te^{2t}, y_3 = -2e^t + 4e^{2t}$
12. $y_1 = c_1e^{2t}, y_2 = c_2e^t, y_3 = (c_2t + c_3)e^t$
13. $y_1 = 3e^{-t} - 3e^{-2t} + e^{-3t}, y_2 = -3e^{-t} + 6e^{-2t} - 3e^{-3t}, y_3 = 3e^{-t} - 12e^{-2t} + 9e^{-3t}$
14. $y_1 = e^{5t} + 7e^{-3t}, y_2 = 2e^{5t} - 2e^{-3t}, y_3 = e^{-5t} + e^{-3t}$
15. $y_1 = \frac{1}{2}e^t - e^{2t} + \frac{1}{2}e^{3t}, y_2 = -e^{2t} + e^{3t}, y_3 = e^{3t}, y_4 = e^{4t}$
16. $y_1 = 2e^{2t} - 1, y_2 = 2e^{2t} - t - 2, y_3 = 2e^{2t}, y_4 = e^{2t}$

9.16 习题

2. (c) $y_1 = (b-1)e^{\lambda} + 2(c+1-b)xe^x + 1, y_2 = ce^x + 2(c+1-b)xe^x$
4. $y_1 = -\frac{1}{3}e^t - \frac{1}{6}e^{4t} + \frac{1}{2}e^{2t}, y_2 = \frac{2}{3}e^t - \frac{1}{6}e^{4t} + \frac{1}{2}e^{2t}$
5. (a) $B_0 = B, B_1 = AB, B_2 = \frac{1}{2!}A^2B, \cdots, B_m = \frac{1}{m!}A^mB$

   (b) $B = -m!(A^{-1})^{m+1}C$
6. (a) $Y(t) = (I + tA + \frac{1}{2}t^2A^2 + \frac{1}{6}t^3A^3)B$, 此处 $B = -6A^{-1}C = -\frac{3}{128}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. 这给出特解 $y_1 = y_2 =$

$-\frac{3}{128}-\frac{3}{32}t-\frac{3}{16}t^2-\frac{1}{4}t^3$

(b) $y_1=y_2=-\frac{3}{128}-\frac{3}{32}t-\frac{3}{16}t^2-\frac{1}{4}t^3+\frac{131}{128}e^{4t}$

7. $\boldsymbol{E}=\boldsymbol{B}, \boldsymbol{F}=\frac{1}{\alpha}(\boldsymbol{AB}+\boldsymbol{C})$
8. (a) $y_1=-\cos 2t-\frac{1}{2}\sin 2t, y_2=-\frac{1}{2}\sin 2t$

   (b) $y_1=2\cosh 2t+\frac{5}{2}\sinh 2t-\cos 2t-\frac{1}{2}\sin 2t, y_2=\cosh 2t+\frac{1}{2}\sinh 2t-\frac{1}{2}\sin 2t$

9.19 习题

4. (c) $Y(x)=e^x e^{\frac{1}{2}x^2\boldsymbol{A}}\boldsymbol{B}$
5. 若 $\boldsymbol{A}(x)=\sum_{k=0}^{\infty}x^k\boldsymbol{A}_k$，则 $\boldsymbol{Y}(x)=\boldsymbol{B}+x\boldsymbol{C}+\sum_{k=2}^{\infty}x^k\boldsymbol{B}_k$，此处对于所有 $k\geqslant 0$，都有 $(k+2)(k+1)\boldsymbol{B}_{k+2}=\sum_{r=0}^{k}\boldsymbol{A}_r\boldsymbol{B}_{k-r}$

10.6 习题

1. (a) $Y(x)=e^x$

   (b) 若 n 为奇数，则 $Y_n(x)=2e^x-\sum_{k=0}^{n}\frac{x^k}{k!}$；若 n 为偶数，则 $Y_n(x)=\sum_{k=0}^{n}\frac{x^k}{k!}$
2. $Y_3(x)=\frac{x^2}{2}+\frac{x^5}{20}+\frac{x^8}{160}+\frac{x^{11}}{4400}$
3. $Y_3(x)=x+\frac{x^4}{4}+\frac{x^7}{14}+\frac{x^{10}}{160}$
4. $Y_3(x)=\frac{x^2}{3}+\frac{4x^7}{63}+\frac{8x^9}{405}+\frac{184x^{11}}{51975}+\frac{4x^{13}}{12285}+\frac{x^{15}}{59535}$
5. (a) $Y_n(x)=1+\sum_{k=1}^{n}\frac{x^k}{k!}$

   (b) $Y(x)=e^x$
6. (a) $Y_2(x)=1+x+x^2+\frac{2x^3}{3}+\frac{x^4}{6}+\frac{2x^5}{15}+\frac{x^7}{63}$

   (b) $M=2; c=\frac{1}{2}$

   (c) $Y(x)=1+x+x^2+\frac{4x^3}{3}+\frac{7x^4}{6}+\frac{6x^5}{5}+\cdots$
7. (a) $Y_4(x)=x+\frac{x^3}{3}+\frac{2x^5}{15}+\frac{17x^7}{315}+\frac{38x^9}{2835}+\frac{134x^{11}}{51975}+\frac{4x^{13}}{12285}+\frac{x^{15}}{59535}$

   (d) $Y(x)=\tan x=x+\frac{x^3}{3}+\frac{2x^5}{15}+\frac{17x^7}{315}+\frac{62x^9}{2835}+\cdots, |x|<\frac{\pi}{2}$
9. $Y_3(x)=2+x^2+x^3+\frac{3x^5}{20}+\frac{x^6}{10}; Z_3(x)=3x^2+\frac{3x^4}{4}+\frac{6x^5}{5}+\frac{3x^7}{28}+\frac{3x^8}{40}$
10. $Y_3(x)=5+x+\frac{x^4}{12}+\frac{x^6}{6}+\frac{2x^7}{63}+\frac{x^9}{72}$；

    $Z_3(x)=1+\frac{x^3}{3}+x^5+\frac{2x^6}{9}+\frac{x^8}{8}+\frac{11x^9}{324}+\frac{7x^{11}}{264}$

索　　引

E

F

G

H

J